

Social Dominance and Biology: Investigating Female Hormonal Response to Non-Physical Competition

By

Martin A. Sharp

Submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy

College of Humanities and Social Sciences
The University of Edinburgh

2005



Declaration

I, Martin A. Sharp, declare that, except for citations referenced in the text, the work contained herein is my own.

Contents

Contents 1

Figures and Tables 4

List of Appendices 6

Acknowledgements 7

Abstract 8

CHAPTER 1 9

Preamble 9

1. Societal Roles 9

1.1 Biology-Behaviour 11

1.2 Testosterone 12

1.3 Social Dominance 13

1.4 Status Hierarchies 15

1.5 Dominance Achieving Strategies 17

1.6 Thesis Structure 19

CHAPTER 2 21

Review of Literature 21

2. Introduction 21

2.1 Behavioural Endocrinology 22

2.1.1 Female Testosterone Metabolism and Feedback Mechanisms 22

2.1.2 Cortisol 24

2.2 Animal Models 25

2.2.1 Context 27

2.3 Testosterone and Male Competition 28

2.3.1 Basal vs. Reciprocal Model 29

2.3.2 Anticipatory Effects and Hormone Variability 30

2.3.3 Normal Stress Response 32

2.3.4 Confounding Effects of Exercise 33

2.3.5 Outcome and Cognitive Interpretation 35

2.3.6 Mood 36

2.4 Human Females 37

2.4.1 Exclusion Criteria 41

2.4.2 Menstrual Variability 42

2.5 Data Analysis 42

2.6 Cortisol and Competition 44

2.7 Summary 45

CHAPTER 3 47

Quantification of Salivary Testosterone: Optimising and Validating an In-house Enzyme Linked Immuno-Sorbant Assay 47

3. Introduction 47

3.1 Free vs. Un-bound Testosterone 47

3.2 Blood-Saliva Correlation 49

3.3 Normal Testosterone Ranges 49

3.4 Contamination of Salivary Samples 50

3.5 Collection Methods 52

3.6 Assay Techniques 54

3.1 ELISA Principle 55

3.1.1 Procedure (adopted in our laboratory) 56

3.1.2 Direct vs. In-direct Technique 58

3.1.3 Solvent Extraction 59

Assay Optimisation	61
3.2 Introduction.....	61
3.3 Amount of Conjugate.....	61
3.3.1 Method.....	62
3.3.2 Results and Discussion.....	62
3.4 Amount of Antibody	62
3.4.1 Method	62
3.4.2 Results and Discussion.....	63
3.5 Incubation Temperature	63
3.5.1 Method	63
3.5.2 Results and Discussion.....	64
3.6 Incubation Time	64
3.6.1 Method	64
3.6.2 Results and Discussion.....	64
3.7 Volume of Sample	65
3.7.1 Method	65
3.7.2 Results and Discussion.....	66
3.8 Optimised Testosterone Assay Protocol	66
3.8.1 Plate Preparation	66
3.8.2 ELISA Procedure	67
Assay Validation	68
3.9 Introduction.....	68
3.10 Accuracy (recovery experiments).....	68
3.10.1 Method.....	68
3.10.2 Results and Discussion.....	69
3.11 Imprecision	69
3.11.1 Method	70
3.11.2 Results and Discussion.....	70
3.12 Assay Sensitivity	70
3.12.1 Method	71
3.12.2 Results and Discussion.....	71
3.13 Specificity Studies	71
3.13.1 Method	72
3.13.2 Results and Discussion.....	72
3.14 Comparison between Commercial Salivary Testosterone Kit and 'in-house' Assay	73
3.14.1 Method	73
3.14.2 Results and Discussion.....	73
3.15 Summary	75
CHAPTER 4	77
Circadian Activity, Episodic Fluctuation, and Testosterone Baselines	77
4. Introduction.....	77
4.1 Baselines.....	78
4.2 Circadian Studies	80
Study One: Circadian Activity in Female Salivary Testosterone	84
4.1 Method.....	84
4.1.1 Participants	84
4.1.2 Procedure.....	84
4.1.3 Salivary Collection.....	85
4.1.4 Hormone Determination.....	86
4.1.5 Hormone Data Analysis	86
4.2 Results.....	87
Study Two: Reliability of Testosterone on Non-Consecutive Days	90
4.3 Introduction.....	90
4.4 Method.....	90

4.4.1 Participants	90
4.4.2 Procedures	91
4.4.3 Salivary Collection and Hormonal Analysis	91
4.5 Results	91
4.6 Discussion	92
4.6.1 Sampling Times	94
4.6.2 Sources of Error	94
4.6.3 Mechanisms	95
4.7 Summary	96
CHAPTER 5	97
Female Hormonal Response to Non-Physical Competition	97
5. Introduction	97
5.1 Method	101
5.1.1 Participants	101
5.1.2 Study Design	102
5.1.3 Procedure	103
5.1.4 Procedure for Collection of Salivary Samples	104
5.1.5 Hormone Determination	105
5.1.6 Analysis of Hormonal Data	106
5.1.7 Attitudinal Evaluation and Mood Measurements	107
5.2 Results	108
5.2.1 Baseline and Pre-Competition Testosterone	108
5.2.2 Baseline and Pre-Competition Cortisol	109
5.2.3 Post-Competition Testosterone	111
5.2.4 Post-Competition Cortisol	114
5.2.5 Cortisol/Testosterone	116
5.2.6 Testosterone, Outcome and Mood	116
5.3 Discussion	117
5.3.1 Baseline and Pre-Competition Testosterone	118
5.3.2 Baseline and Pre-Competition Cortisol	120
5.3.3 Post-Competition Testosterone	120
5.3.4 Post-Competition Cortisol	122
5.4 General Discussion	123
5.4.1 Psychological Variables	123
5.4.2 The Importance of Context for the Testosterone-Dominance Relationship	125
5.5 Limitations	126
5.5.1 Baselines	126
5.5.2 Mood Evaluations	128
5.5.3 Outcome	128
5.6 Summary	129
CHAPTER 6	131
Epilogue	131
6. Introduction	131
6.1 Chapter Summaries and Future Research Directions	131
6.2 Conclusion	136
References	138

Figures and Tables

Figures

<i>Fig. 1</i>	Bio-synthesis of Testosterone in Women	24
<i>Fig. 2</i>	Components of whole saliva	59
<i>Fig. 3</i>	Standard Curves Produced by Varying the Amount of Conjugate	62
<i>Fig. 4</i>	Effect of Antibody Concentrations on Testosterone Standard Curve	63
<i>Fig. 5</i>	Standard Curves Following Incubation at Varying Temperatures	64
<i>Fig. 6</i>	Standard Curves Produced at Differing Incubation Times	65
<i>Fig. 7</i>	Testosterone ELISA Standard Curves of Varying Volumes of Sample	66
<i>Fig. 8</i>	Correlations between In-house ELISA and Commercially Available Assay.....	74
<i>Fig. 9</i>	Correlation between In-house ELISA and Commercially Available Assay (non-alkaline samples).....	75
<i>Fig. 10</i>	Raw Testosterone Data from 3 Subjects.....	87
<i>Fig. 11</i>	Circadian Female Salivary Testosterone (Mean \pm SEM).....	88
<i>Fig. 12</i>	Percentage Change (\pm SEM) in Salivary Testosterone from Mean Circadian Levels	89
<i>Fig. 13</i>	Percentage Change (Mean \pm SD) in Raw Salivary Testosterone from 9am Levels	89
<i>Fig. 14</i>	Normalised Testosterone Across Two Non-Consecutive Days (Mean \pm SEM).....	92
<i>Fig. 15</i>	Normalised Testosterone at Baseline and Pre-Competition Phase (Mean \pm SEM)	108
<i>Fig. 16</i>	Percentage Change from Mean on Normalised Testosterone Data at Baseline and Pre-Competition Phase (Mean \pm SEM)	109
<i>Fig. 17</i>	Normalised Cortisol (Mean \pm SEM) at Baseline and Pre-Competition Phase	110

<i>Fig. 18</i>	Percentage Change from Mean (Mean \pm SEM) for Normalised Baseline and Pre-Competition Cortisol	110
<i>Fig. 19</i>	Normalised Post-Competition Testosterone by Outcome (Mean \pm SEM).....	111
<i>Fig. 20</i>	Mean Normalised Post-Competition Testosterone by Outcome at 30 mins and 1 hr	112
<i>Fig. 21</i>	Post-Competition Testosterone Percentage Change from Mean by Outcome (Mean \pm SEM)	113
<i>Fig. 22</i>	Post-Competition Cortisol Levels by Outcome (Mean \pm SEM)	114
<i>Fig. 23</i>	Percentage Change in Post-Competition Cortisol Levels by Outcome (Mean \pm SEM)	115
<i>Fig. 24</i>	Normalised Cortisol and Testosterone for Winners (Mean \pm SEM)	116
Tables		
<i>Table 1</i>	Range for Female Salivary Testosterone in Hormone-Competition Studies	50
<i>Table 2</i>	Methods of Salivary Stimulation in Hormone-Competition Studies	53
<i>Table 3</i>	Recovery of Known Amounts of Unlabelled Testosterone Added to Saliva Pool	69
<i>Table 4</i>	Cross Reactivity between Testosterone And Related Steroid Hormones	72
<i>Table 5</i>	Testosterone Baselines: Sampling Times in Hormone-Competition Studies	79
<i>Table 6</i>	Individual Variability in Testosterone Levels from Circadian Mean	89
<i>Table 7</i>	Reliability of Salivary Testosterone across Two Non-consecutive Days	91
<i>Table 8</i>	POMS Data at Baseline, Pre, and Post-Competition (Mean \pm SEM)	117

List of Appendices

Appendix A	Chemical Reagents and Consumables	159
Appendix B	Circadian Activity Exclusion Criteria	160
Appendix C	Circadian Activity Informed Consent	161
Appendix D	Circadian Activity Detailed Instructions	162
Appendix E	Hormonal Response to Competition informed Consent	165
Appendix F	Hormonal Response to Competition Detailed Instructions	166

Acknowledgements

Thanks are tendered to Andy McKinlay in gratitude for his insight, personal integrity, assistance and friendship.

Thank you also to Emad Al-Dujaili, for his warm friendship and unstinting contribution to my laboratory tuition and education.

For Charis Xinari, who has supported me unerringly, whose own studies on gender politics proved invaluable for helping to develop the framework of this doctoral thesis, and whom I love dearly.

Abstract

Within the broad framework of evolutionary theory it is possible to develop a sustainable foundation for social dominance-hormone relationships in women. However, whilst providing an important contribution towards understanding the role of biology in social dominance, hormone-competition interactions have received considerably less attention in females than in males. Consequently, the thesis explores the relationship between salivary testosterone (T), cortisol (F), and non-physical competition in women.

In order to address widely acknowledged difficulties with determining levels of female T, particularly the biologically active 'free' fraction as measured in saliva, a highly sensitive 'in-house' enzyme-linked immunosorbant assay (ELISA) was optimised and validated. Assay sensitivity was 0.5pg/mL; intra and inter-assay coefficients were 2.1% and 6.7% respectively; recovery ranges were between 96-105%; cross reactivity with related compounds was minimal – i.e. androstendione <8%. A modest correlation with a commercially available T assay kit is discussed in relation to the limitations inherent in direct enzyme assays. The ability to measure salivary free T at levels below those easily available to users of commercial kits is fundamental to the work in this thesis.

In the absence of reliable information on the daily profile of female free T, bio-behavioural studies have tended to formulate salivary sampling strategies taken directly from research involving male subjects. However, this approach may be inappropriate. Accordingly, by determining a comprehensive picture of the circadian activity of salivary T in 36 healthy female subjects, it was possible to demonstrate that T follows a circadian rhythm, the relative levels of which differ over two non-consecutive days. Moreover, throughout the course of the day T levels were highly variable, with episodic fluctuation of individual data points exceeding 83% of 9am levels. These findings highlight the necessity of collecting multiple samples in bio-behavioural research involving T and women.

Incorporating methodological refinements in both the measurement of T and sampling protocol, an experimental study sought to examine the dynamic relationship between T, F, and non-physical dyadic competition. Twenty-two females (ages 19–24 years) competed in a knockout tournament involving the wood-block game 'Jenga'. They collected comprehensive salivary samples for baseline, pre- and post-competition phases. Subjects additionally reported mood states and answered questions concerning their participation in the competition. Whilst the comprehensive hormonal data resist easy interpretation, compared against baseline levels, pre-comp T and F appeared unresponsive in anticipation of competition. However, at 3 hrs post-competition, T levels rose 35% in winners and fell 4% in losers. F levels, conversely, fell in both losers and winners, although in winners this drop was much more pronounced. These results suggest that, similar to males, women experience a dynamic endocrine response following competition. However, the extent to which T is implicated in female social dominance is likely to be extremely small. The findings further illustrate how choice of competitive task and hormonal sampling regimens used in previous studies may, to a large extent, account for the equivocal findings in the literature. Additional research is required to ascertain if this pattern of endocrine response holds under a wider variety of competitive situations and also to explore more fully the psychological processes mediating hormonal responses to competition.

1. Societal Roles

Female oppression is inscribed by patriarchy's attempt to use women's biology as a justification for social inequality between men and everyone else who is *other* than men (Xinari, 2004). In *Myths of Gender: Biological Theories about Women and Men*, her discussion of biologically based sex differences, Anne Fausto-Sterling sees women's biology as a convenient self-justification for their secondary social position which also ensures that the 'facts' of biology become central in the narrative of women's oppression (Fausto-Sterling, 1992).

Simone de Beauvoir's *The Second Sex* (1949) opens with a detailed discussion of 'The Data of Biology'. To 'be a woman', she says in her introduction, is a problematic position in which to be placed: on the one hand, feminism (dedicated to leading the struggle for social equality between the sexes) considers the term 'woman' as carrying negative and derogatory connotations and is uncomfortable with its use; on the other hand, misogynist views, in trying to secure woman's inferior status in society, call upon women to remain 'women' and locate femininity (the social aspects of gender) as sourcing from what they consider to be 'the facts of biology'. Beauvoir does not deny the facts of biology; her detailed account of them, as she admits, highlights their importance and their contribution to how the category 'woman' has been formed. She is, however, interested in discussing how they have come to create for woman a fixed position in society. Beauvoir aims to explore how biological differences between the sexes have come to fix an inevitable position for women as the 'second sex'. Hence, a question of fundamental importance for feminist scholars, psychologists, and social theorists, concerns whether women occupy their place in society by dint of their biology; does it make them, in some way, inherently inferior. Or, does this inequality have socially constructed roots (Fausto-Sterling, 2000; Kemper, 1990)?

The major ideology of male-female relations has specified male ascendancy and female subordination; social arrangements of male dominance precluding pre-eminence for women. In addition, there were, and to a great extent there still remain, prescriptive models of the normal, natural, and desirable personality of women (Kemper, 1990). In one illuminating study, Broverman et al. (1970) examined the assumptions of psychotherapists about what constituted a healthy female personality. The results are instructive and, sadly, typical of a particular kind of widespread misogyny that has existed within the sphere of science. Women were thought of as being dependent, gentle, subjective, empathic, and, importantly in the context of this thesis, non-competitive. Whilst women may indeed possess these traits, it is significant that they were not thought to possess traits of dominance, ascendancy, or other power or status oriented characteristics that are attributed to males.

Although these findings emanate from a study conducted only thirty years ago, as a view of women they appear strangely old-fashioned. Noting that these traits are not in accord with a growing body of evidence Kemper (1990) states, 'Across the spectrum of the social sciences, data from more recent studies suggest that females are not essentially retiring, unaggressive, or lacking in motives and psychological needs for power and dominance. Thus, while successful ideological socialisations may persuade many women that this is true of themselves, it is not biologically true' (p.138). During recent decades, female aggression has been studied extensively, especially the female propensity for indirect aggression. During their reviews, Eagly and Steffen (1986) and Frodi, Macaulay, and Thome (1977) concluded that when aggressive, females try to cause psychological rather than physical harm to their enemies. Hence, although less physically aggressive than males, women have found other ways to aggress against their enemies. Hrdy (1999), having studied a number of animal species, makes the case as firmly as anyone by suggesting it would be a mistake to think that the female of many species could afford to be unaggressive, either in defence of territory or her offspring. Nor, as the evidence increasingly shows, is she unassertive in sexual relations.

Cross-culturally, the gender difference in the propensity for physical aggression appears marked, with males committing significantly more violent crimes than females (Eagly & Steffen, 1986; Frodi et al. 1977). However, Muncer et al. (2002) report that,

in the late 1990's, 17% of all violent crime committed by people aged under 21 was by females.

1.1 Biology-Behaviour

Much of the early science of human social behaviour was conceived against the backdrop of a dichotomy between biology and culture, with the two being set in binary opposition to one another. Whereas biology has been viewed as intractable and fixed, culture has been afforded a much more malleable role. With rapid and astonishing advances in the study of human genetics, however, many scientists now interpret this dichotomy as misleading and false (c.f. Ridley, 2003). In fact, the two are highly interdependent upon one another, to the extent that Fausto-Sterling (1992) argues that questions concerning the relative contribution of either are essentially meaningless. As Goldstein notes, 'Biology provides diverse potentials, and cultures limit, select, and channel them. Furthermore, culture directly influences the expression of genes and hence our biology' (2001, p.251). Nevertheless, amongst a wide range of influential sociologists, anthropologists, social psychologists, and feminist scholars, the study of biological bases of human behaviour has been, and remains, strongly contested (Kemper, 1990). The reasons for this oppositional stance can be viewed as emanating from early simplistic arguments leading in the direction of biological necessity as an explanation of present day gender relations (Goldstein, *ibid.*). Nevertheless, as Dabbs (2000) notes, 'People are biological creatures who live in social settings, and any good explanation of how testosterone works must be a biosocial one' (p.196).

It is widely believed that hormones must play some role in human male aggression; even though the term aggression is invariably merged with a loose constellation of terms such as dominance, hostility, predation, competition, and initiative (c.f. Brinkerhoff & Booth, 1984). Whilst the interplay between the endocrine system and competitive or dominating behaviour is ferociously complex '...involving both hormonal effects on behavior and experiential effects on endocrine function' (Gladue, Boechler, & McCaul, 1989, p.409) it is not inconsequential to consider the role of hormones on human female behaviour; lying as they do at the interface between genetics and culture (Rogers, 1999).

1.2 Testosterone

Testosterone (hereafter T) is one hormone in a group of compounds called androgens. It exerts powerful physiological effects and is associated with, amongst other things, the development of male secondary sex characteristics. For example, it promotes growth of the reproductive tract, development of the prostate, and the production of body hair. As well as these androgenic or masculinising effects, T also drives anabolic, or tissue-building, changes. These include thickening of the vocal chords, growth spurts, development of sexual libido, and an increase in strength and muscle bulk (Mooradian, Morely, & Korenman, 1987). Whilst estrogens and progesterone are the so-called female sex-hormones, T has been termed the male sex-hormone. In recounting the fascinating story of how these hormones were named, Fausto-Sterling (2000) has persuasively argued that the label of sex-hormones is highly misleading given the myriad other functions they serve. However, with the emphasis firmly on the sex aspect of their function research agendas have been subsequently channelled, acting as a crutch for patriarchal ideologies which have proved almost impossible to overcome. For example, Björqvist et al. (1994) identify that, whilst findings from a study involving T and aggression in newborn female mice was well publicised, a further study by the same authors (Edwards and Herndon, 1970), in which it was demonstrated that oestrogen also increased fighting behaviour, 'has gone almost unnoticed' (p.19). The reasons they propose for this centre on the difficulties associated with explaining how a supposedly female sex-hormone increases aggressive behaviour when women were not thought to be aggressive.

In fact, whilst men produce their own oestrogen, T is also part of the female endocrinological profile; it is synthesised in the ovaries and adrenal cortex but also around 50% is metabolised through peripheral conversion of androgen precursors (Somboonporn & Davis, 2004). In this regard, there is a growing wave of publications concerning the importance of androgens to women's health (i.e. Davis & Tran, 2001). For example, T is thought to contribute to libido and is important in maintaining muscle mass and bone strength. Moreover, Simpson (2002) tells us, 'Testosterone circulates at concentrations that are an order of magnitude greater than those of estradiol in the postmenopausal women. An obvious implication of this realisation is that androgens have an important role to play in female physiology' (p.S6).

However, since rightly or wrongly, T is regarded as the sex hormone that imbues males with their predominantly dominant, competitive, and aggressive characteristics, it has been investigated more in males, just as oestrogen has been investigated more in females. Kemper (1990) suggests, however, that whilst little is currently known about the effects of T on female social behaviour, the data that does exist suggests tremendously important social effects of T levels in females, parallel in many ways to the effects in males. In as much as gender differences in biological structure or function have been used to support the notion of women as somehow biologically inferior to men, findings associated with the endocrinological foundation of competition in males are provocative. Thus, it is of interest to know whether this hormone operates in the same manner in females, whether it is responsive to the same social stimuli and incentives and, if so, whether it produces the same social effects as in males. Caution is urged, however. What emerges from the puzzling and inconsistent research is that interpretation of T appears too often caught in a confusing web of social expectations, ideological positions, and gender stereotypes. As van den Berhge (1974) notes: 'The more a subject becomes ideologically laden, the greater the disagreement is in interpreting the facts' (p.681).

1.3 Social Dominance

A significant body of research has sought to elucidate the relationship between T and aggression, particularly in males. However, Archer (1988) suggests that in addressing this topic researchers are confronted with seemingly insurmountable ideological, methodological, and semantic traps. Despite considerable attempts at elucidation, the role played by T in human aggression is still fervently contested, with Archer (1998) calling the results 'confusing and contradictory' (p.363). In formulating his biosocial model of status, Mazur (1985) proposed 'It may be the case that circulating testosterone is related primarily to dominance behaviour among men and not to aggression except in situations where dominance happens to be asserted aggressively' (p.382). Indeed, although a number of scientists strongly refute the claim, Dabbs (1998) describes the relationship between T and aggression as '...an idea whose time has come and gone' (p. 370).

No longer regarded as a unified concept, aggression is thought to comprise distinct components, often with differing underlying hormonal, genetic, and social control mechanisms. In an attempt to reflect some of this complexity, Moyer (1968) proposed seven types of aggression: predatory, inter-male, fear-induced, irritable, maternal and sex-related. Although this typology has since been criticised (e.g. Archer, 1988) and amended, the difficulties of definition still remain. In this regard, Brinkerhoff and Booth (1984) state, 'The term aggression refers to a loose collection of activities and attributes that are often unrelated to one another. Predation, initiative, dominance, territorial behaviour and hostility are among the ideas incorporated into studies of aggression' (p.159). The primary point here concerns the essential challenge of definition, with Archer (1998) suggesting there is no reason to think that the concept of dominance is likely to prove any less problematic in this regard.

In a commentary on Mazur and Booth's target article *Dominance and Testosterone in Human Males*, Hines (1998) suggested the authors' use of indirect evidence included behaviours that are not typical of dominant individuals, nor do they produce dominance. Thus, we are immediately alerted to the possibility, indeed probability, that conceptualisation as to what constitutes not only dominance but also evidence of dominant behaviour is going to vary significantly between researchers. And indeed, dominance has been conceived of in a number of different ways (Ellyson & Dovidio, 1985). Social dominance has been alternately described as: '...an apparent intent to achieve or maintain a status advantage, such as power, prestige, or valued prerogatives or resources' (Gladue, Boechler, and McCaul, 1989, p.410). 'The core of dominance can be identified as acting overtly so as to change the views or actions of another' (Fiske, 1971, p.98). Mazur (1985) employed the following definition, 'An individual will be said to act dominantly if its apparent intent is to achieve or maintain high status over a conspecific' (p.382). Finally, Cashdan (1995) suggested that dominance rank '...is a reflection of the outcome of competition' (p.355). Although just a selection of definitions, there is a divergence of thought here. A number of authors clearly envisage dominance as a personality trait (i.e. Moskowitz, 1993; Mudrack, 1993; Weisfeld, 1994), whilst others are at pains to point out that social dominance is actually a descriptive term applied where there is inequality in power relations or competitive abilities (i.e. Archer, 1994; Hawley, 1999; Tinbergen, 1953). This debate is nowhere near a resolution and, unfortunately, a thorough examination of the main arguments in

this important arena would lead too far away from the main aims of the thesis. At this juncture, identification of the schism that exists will have to suffice.

1.4 Status Hierarchies

Mazur (1973) describes the multi-dimensionality inherent in the organisation and maintenance of small group structure in primates. He suggests that 'Kinship, courtship, and consort linkages all provide additional structure, sometimes relating to status hierarchy and sometimes independent of it. Interaction...may be shaped by these...links more than by...status relationships' (p.380). However, as the behaviour of social species evolved in the context of the social group, in both animals and humans, dominance is nevertheless an important dimension of social interaction, even if it does not explain all social behaviour (Brewer & Caporael, 1990; Cosmides & Tooby, 1987; Mithin, 1996; Trivers, 1971). Indeed, a number of authors note that, through status differentiation, the emergence of dominance hierarchies in small groups is documented across every culture throughout history (Dunbar, 1988; Mazur, *ibid.*). In considering why this might be, it is of note that animals living in groups are likely to benefit from reduced predation risk, easier access to food, increased productivity, availability of mates, or any combination of these factors (Camazine et al., 2001). However, where there is increased competition over critical resources (as there invariably is), conflict can and does arise from group living (Darwin, 1859). In many species of animals, when several un-related individuals are placed together in a group under the standard experimental resident-intruder paradigm, they engage in contests for dominance. These interactions take various forms; physically violent fights, aggressive displays, mock-fights (ones that do not incur any serious injury), and some are limited to the passive recognition of a dominant and a subordinate. Often, there are multiple interactions, taking place over periods ranging from hours to weeks. Depending upon factors such as group size and species of animal, contests will vary in frequency, becoming less frequent and eventually being replaced with more or less stable dominance-subordination relations amongst all group members (Chase, 1974; Wilson, 1975). When such a structure arises it organises the group in such a way that conflicts do not completely offset the advantages of group living (Richard & Schulman, 1982).

Whilst several authors have investigated the formation of dominance hierarchies, little research has focused on gender differences (Schmidt Mast, 2001), even though the stereotypical view of men being inclined to form hierarchical dominance structures while women build egalitarian structures is widely accepted (de Waal, 1986; Gilligan, 1982; Moskowitz, Suh, & Desaulniers, 1994). In this respect, Mazur (1973) noted, even in primates 'It is often easier to identify a male ranking than a female ranking' (p.518). These gender differences, where women are considered to be less overtly competitive than males, have been explained in terms of the socialisation into cultural norms (Miner & Longino, 1987; Tracy, 1991). Given the suggestion that the connection between social dominance and androgens is of central importance in the evolution of male social behaviour, it is legitimate to enquire after the existence of a female equivalent. And, although contrary to the findings of earlier studies, Cashdan (2004) has argued persuasively that, whilst female hierarchical structures may take longer to become stable, men and women may actually be equally concerned with forming these structures within a group setting.

Considered within a broad framework of evolutionary theory, the implications of caring for offspring (parental investment), which in humans at least falls disproportionately to women, ensures the potential cost of aggressive competition in females is high (Trivers, 1972). Hence, it is possible to contend that if hormones are implicated, even partially, in the motivation for and maintenance of competitive behaviours – where these behaviours are likely to lead to a high risk of physical injury - then women may well not have evolved an endocrine response to competition. In this respect, Campbell et al. (1998) argued that T is not instrumental in female behaviour because women restrict their competition to things that have no status implications. This interpretation is at odds with Hrdy (1981), however, who notes that competitiveness is useful for females too, especially in vying for the support of males, protecting territory, and in protecting and providing for offspring; it is difficult to imagine there is no status attached to these events.

Contrary to early accounts of evolutionary fitness in females, it appears that women do possess high reproductive variability and the benefits of gaining access to resources are high. Dominant females have more offspring, their female offspring reach reproductive age quicker, and have greater infant survival (Ellis, 1995; Grant, 1994).

And, within certain species of primates, male ranking is largely dependent upon the mothers' position of dominance with the group (Grant, *ibid.*; Koford, 1963; Mazur, 1973). Moreover, hierarchies or social structure need not be exclusively patriarchal. Among certain animals there exist matriarchal dominance hierarchies, such as those found in Hyenas or Bonobo chimpanzees. Consequently, the view that males are naturally more inclined to form dominance hierarchies than women stands in stark contrast to the paucity of research and contradicts a number of theoretical positions (i.e. Mazur's biosocial theory of status in face to face interactions (1985). Moreover, Grant and France (2001) and the theorist Kemper (1990) suggest that we are too ready to dismiss the possibility that a female's endocrine response to competition may be similar to that of men.

1.5 Dominance Achieving Strategies

Whilst physical aggression is certainly one strategy employed by males as a means of achieving a position of dominance within a hierarchy, it is not by any means the only one. Nor, in humans, is it arguably even the most important strategy. Indeed, as Mazur and Booth (1998) observe, 'Much interpersonal behaviour is overtly or subtly concerned with managing dominance or subordination without causing physical harm' (p.353). Cashdan (1995) says about dominance, 'Dominance rank is a reflection of the outcome of competition, and assertiveness and toughness might be expected to enhance success in competition' (p.355). From a divergent stance, Dabbs (1998) suggested that whilst T is related to dominance, dominance need not be competitive but, '...can arise from strong personal characteristics that produce admiration and deference in others' (p.370). Dabbs (*ibid.*) further suggests that while dominance is usually construed as competitive, where one gains at the expense of another, '...there are many ways in which people dominate or control others and sometimes this is to the advantage of all concerned' (p.371).

Amongst animals as well as humans, not only is there competition (aggressive or otherwise) but there is also considerable coalition behaviour (van Hoof & van Schaik, 1992). As de Waal (2001) observes 'Individual abilities do play a role in achieving high status, but the abilities involved are often distinctly social, such as diplomacy and a

talent for building lasting partnerships' (p.300). Consequently, through a series of often complex cognitive processes, ranks may be determined co-operatively, in which case dominance might more accurately be considered as eminence (Kemper, 1990). Closely related to the idea of eminence is the proposition that hierarchies are a function of expectations about task performance, where those individuals considered most competent to assist the group in optimal performance assume a position of dominance (Ridgeway & Berger, 1986; Ridgeway et al. 1985). Clearly, these strategies are assisted by the human facility for verbal fluency. Thus, for males at least, there are differing means of achieving a position of high status and consequently securing access to valued resources, such as a mate in oestrus. And so, if T is implicated in dominance related behaviour, even in a general or permissive way, it has to be shown to influence behaviours other than aggression.

Campbell et al. (2001) discussed a number of aspects of evolutionary theory that underpin reasons why women may be less competitive than males suggesting that rather than being related to aggression exclusively fear thresholds play an important role in the disparity between genders in relation to aggressive behaviour. There are good reasons, therefore, to expect dominance seeking in women to be as variable as it is in men and there is a good deal of evidence that whilst women employ more indirect methods of achieving dominance such as tactics of exclusion and reputational attacks (Hess, 2004) there has recently been an upsurge in levels of often extreme physical violence amongst adolescent western females (Campbell, 1999).

It may be that many of the non-competitive aspects of dominance may well be related to T and it would be short-sighted to focus only on competition. That said, competition has an important part to play in achieving a dominant position within a societal structure. Arguably, in Western human societies this competition most often takes the form of non-physical encounters. However, this is only true in certain strata of society. There are still abundant instances where physical prowess is highly regarded, such as the honour subcultures described by Cohen et al. (1996). With specific regard to women, a small number of studies have claimed that physical aggression is higher in non-white women who are poorly educated and of low socio-economic status (Brown, 1998). In this thesis I investigate the endocrine response to non-physical competition; acknowledging that this can represent only a tiny although important fraction of what is

thought of as dominance. In talking about the relationship between T and social dominance Mazur and Booth note, however, that ‘...competition studies...are central to our analysis of the reciprocal model’ (p.387).

1.6 Thesis Structure

Whilst this thesis is centrally concerned with exploring the relationship between hormones and female non-physical dyadic competition, significant effort is expended on ensuring accurate measurement of salivary un-bound T. Structurally then, the thesis can be thought of as addressing two distinct although associated issues. Following this introductory chapter, in which evidence of a sustainable link between hormones and social dominance in women is presented, chapter 2 - a review of the literature - frames fundamental questions about the relationship between hormones and competition in females. An initial brief examination of the literature on competition and androgens in animals is used to identify methodological pitfalls for human studies. The chapter proceeds to summarize the findings from male hormone-competition studies’, identifying a number of methodological confounds that limit the reliability of results. Finally, the two published studies involving female participants are considered.

Measurement of the free, biologically active, component of T is a critical element in determining the nature of hormone-competition interactions. This, more than any other methodological aspect comes up against the limits of technology. It is possible to accurately measure free T at the low levels found in females, but not easily and not straightforwardly. Subsequently, chapter 3 is concerned with the optimization and subsequent validation of an ‘in-house’ enzyme-linked immunosorbant assay. In order to address concerns over measurement in context this chapter not only highlights the challenges faced in the measurement of salivary T in women, but also reviews the relevant literature. For example, issues surrounding the use of saliva as a diagnostic fluid and the enzyme linked immunosorbant technique as a means of measuring female salivary T are introduced.

In attempting to understand the dynamism between hormones and competition, studies have tended to determine T levels with only a limited number of measurements;

several studies collecting one sample pre and one further sample post-competition. However, there is little detailed and accurate information on the circadian activity of female free T and any circadian activity may confound the results from studies employing this methodology. Consequently, chapter 4 is an attempt at ascertaining the validity of single time-point sampling protocols through an investigation into circadian activity of salivary free T in women.

Incorporating the methodological advancements emanating from the findings of the circadian studies, and utilising the previously optimised assay, chapter 5 seeks to investigate directly the extent to which there is evidence of a dynamic relationship between female salivary T and cortisol (F) in relation to non-physical competition. The final chapter provides a summary of individual chapters and suggests future research strategies for furthering our understanding of the role of hormones in female social dominance.

2. Introduction

There is growing evidence to suggest the existence of a dynamic relationship between the endocrine system and behavioural competition¹ in humans. Implicated in such behaviours and psychological states as dominance, anger, aggression, hostility, and assertiveness, the role of testosterone (T), more than any other hormone, has been the subject of intense speculation and enquiry. Findings from a number of studies reveal a clearly identifiable endocrine response to competitive encounters, with T levels rising prior to competition and subsequently remaining elevated following victory. Where competitors are defeated, T levels fall. In proposing the bio-social model of status in face-to-face primate groups Mazur (1985) hypothesised that this divergence in T levels may underpin status ranks in formation of dominance hierarchies. The theorist Kemper (1990) also carefully considered this possibility in the wider context of human society.

The literature addressing hormone-competition interactions has made use of various experimental protocols and employed differing approaches to statistical analyses; combined, this ensures that simple, straightforward, interpretation of the literature is problematic. Nevertheless, it is the aim of this review chapter to describe what is currently known about the interaction between hormones and competition; subsequently identifying and exploring some of the methodological and conceptual weaknesses which limit interpretations that can usefully be drawn from the results of these studies. In order to facilitate this aim, the chapter is structured in the following manner. Firstly, an overview about the nature and role of T in humans emphasises where important sex differences exist; in sites of androgen production for example.

Secondly, serving as the theoretical foundation for considering parallel effects in humans, the nature of interactions between hormones and dominance in animals are briefly discussed. Following elucidation of the relationship in non-human primates, findings from studies examining the aspects of the same relationship in humans are presented and evaluated. As studies investigating the effects of competition on the endocrine system in women are scarce (currently two published studies), I initially consider findings from studies involving male participants. Following a synopsis of the current findings, various confounding elements are identified and their potential effects on the validity of findings are addressed in greater detail. Finally, I consider the current state of literature exploring the same relationship in human females. Given the paucity of studies available these are considered in greater depth.

2.1 Behavioural Endocrinology

In response to wide fluctuations in the external environment, the endocrine system helps maintain the steady state of our bodies. Consisting of endocrine glands that secrete hormones, the hormones themselves, and target tissues that respond to the hormones, the endocrine system controls metabolism, growth, reproduction, and assists in adaptation to stress and changes in physical circumstances (Chiras, 1999). It also regulates the concentrations of important substances in the blood, such as glucose, calcium, sodium, potassium, and water. Secreted by various endocrine glands throughout the body, as well as by the neurones in the brain, hormones function as chemical messengers - transported by the blood to target tissues, where they activate changes in physiological activity. Flinn et al. (1998) describe the neuro-endocrine system as ‘...a complex set of mechanisms designed by natural selection to communicate information among cells and tissues’ (p.372).

2.1.1 Female Testosterone Metabolism and Feedback Mechanisms

T is synthesised from pregnenolone, which is formed from cholesterol (Griffen & Wilson, 1992; Hall, 1988). In plasma, T circulates non-specifically bound to proteins

¹ Competition can take many extremely subtle forms at the biological and neuro-endocrine level. Consequently unless otherwise stated all future reference to competition refers to the behavioural manifestation

(albumin) and specifically bound to sex-hormone binding globulin (SHBG), so that only 1-3% of the circulating concentration is 'free' to exert its biological activity. Metabolic pathways for androgens are complex. Partial dissociation of the specific steroid-protein complex may occur and inactive pro-hormones may be converted intra-cellularly to active hormones. Thus, the 'free' and non-specifically bound hormone fraction in plasma (commonly referred to as the bio-available fraction) only partially reflects the hormone available at the cellular level in specific tissues (Vermeulen, Verdonck, & Kaufman, 1999). Nevertheless, there is sound evidence that this 'free' fraction has greater biological significance than the total hormone levels in plasma (Vermeulen & Verdonck, 1972). This issue - how various components of circulating T are transported in the blood - has important implications for the measurement of T, and is subsequently explored more fully in the following chapter.

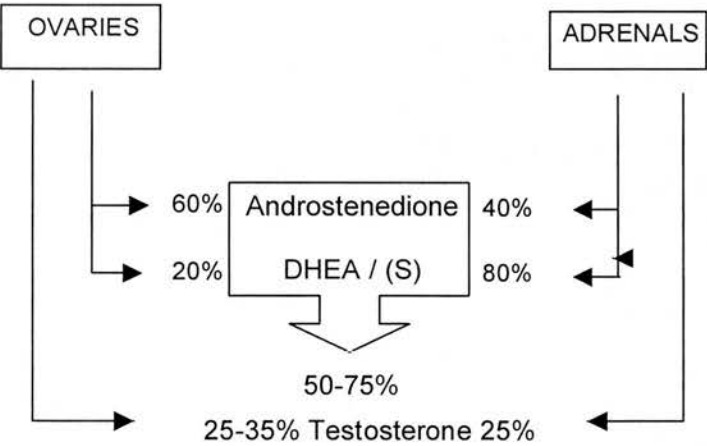
In males, T is synthesised primarily in the leydig cells of the testes (Parker, 1991). The release of T operates on a classic negative feedback system, based on the release of gonadotropin releasing hormone (GnRH) in the hypothalamic region and both luteinizing hormone (LH) and follicle-stimulating hormone (FSH) from the anterior part of the pituitary gland. In women, however, T synthesis is more complex, with T production occurring at several separate sites (Burger, 2002). Figure 1 illustrates the basic bio-synthesis of T in women. Roughly, one quarter of the hormone T is produced in the ovary (ovarian stroma), a quarter is produced in the adrenal gland (adrenal zona reticularis). The ovaries and adrenals additionally produce the androgen precursors: androstenedione and DHEA, which are subsequently sulphated by the enzyme sulphataze into DHEA-S. These androgen precursors yield the remaining 50% of T (Guay, 2001; Longcope, 1986).

T is positioned within a chain of precursors and metabolites. Consequently, there is also a significant amount of inter-conversion among steroid hormones on metabolic pathways. The main T precursor in the pre-menopausal ovary is androstenedione, which is converted primarily to estrone, but which can also be converted to androgens (Guay, 2002). The predominant precursor in the adrenal gland is dehydroepiandrosterone (DHEA) which is mainly excreted in the sulfated form, dehydroepiandrosterone sulphate (DHEA-S) (Parker, 1991). At a cellular level T exerts its effect through partial

conversion to the more powerful metabolite 5 α -dihydrotestosterone (5 α -DHT) or estradiol (Simpson, 2002).

Reflecting the different sites of production, rather than one feedback system, T in females appears to operate under the control of at least two feedback loops. The release of T from the ovaries is regulated, via GnRH of the hypothalamic region, and by FSH and LH of the pituitary gland (Rogers, 1999), whereas production of T from the adrenal cortex is adjusted via the corticotrophin releasing hormone (CRH) from the hypothalamic region and by adrenocorticotrophic hormone (ACTH) in the pituitary gland. It is worth noting that where authors such as Filaire et al. (2001) have suggested that levels of cortisol (F) and T in women may follow similar profiles in respect of competition they neglect to acknowledge the fact that only around one quarter of T is produced at the same site as F. This theme is returned to during the discussion of the findings in chapter 5.

Figure 1 Bio-synthesis of Testosterone in Women (adapted from Somboonporn and Davis, 2004)



2.1.2 Cortisol

In humans, F is a principal hormonal product of the hypothalamic-pituitary-adrenal (HPA) axis (Miller & Tyrell, 1995) and, like T, is also synthesised from cholesterol. In order to maintain vital systems and processes (such as ensuring adequate glucose for the brain) F will sacrifice tissue such as muscle, fat, and protein and is, hence, considered the most important glucocorticoid in humans. F also acts to counteract the effects of

stress and is correspondingly known most commonly as the 'stress' hormone. Although the exact nature of the response is not fully understood it is thought to include cardiovascular, neurological, and anti-inflammatory effects.

There is a wide variation in 'normal' values between individuals (Riad-Fahmy et al., 1982). Levels are in the range of 12-16nmol/L and secretion has an ACTH dependent diurnal rhythm, which can fluctuate widely during each day (Clow et al., 2004; Kudielka et al. 2003; Shimada et al., 1995). Secretion of F is controlled by the HPA axis and it is secreted into the circulation in response to ACTH which is, in turn, secreted by the anterior pituitary in response to hypothalamic corticotrophin releasing factor (CRF). The release of CRF is affected by circadian rhythm, stress, and several other psychosocial factors. Among the many factors contributing to individual variations in HPA activity gender appears to play a complex role (Kirschbaum et al., 1999). Several studies have demonstrated that, in response to psychosocial stress, men display higher F levels than women (Kirschbaum et al., 1999; Smyth et al., 1998). However, the effect of gender on F reactivity under basal conditions remains unclear. As Kirschbaum et al. (1999) note: 'Among the multitude of factors contributing to individual differences in HPA activity, a complex sexual dimorphism seems to exist' (p.154). As with T, the assessment of F in saliva has become a useful alternative to blood testing. Only 5-10% of plasma cortisol is to be found in the 'free' form; the remaining F is bound with protein with only the unbound fraction available to the tissue. The acinar cells lining the salivary glands prevent proteins and protein-bound molecules from entering saliva; therefore the measurement of F in saliva appears to be a reliable method for assessing the bio-available component of F in the body (Kirschbaum & Hellhammer, 1994).

2.2 Animal Models

As Mazur (1973) notes, drawing analogies between species is interesting but may not be especially meaningful '...particularly when the two species are widely separated on the phylogenetic tree' (p.515). Indeed, there are many differences in the nature of T interactions among various species of animal, as well as among various strains of a given species (Chambers, 1998). Thus, no one animal species can legitimately serve as

an exhaustive model for considering the nature of hormone-behaviour interactions in humans (Hrdy, 1999). Indeed Dixon (1980) goes further, suggesting ‘...interspecific variability and differing experimental conditions make it impossible to generalise’ (p.39). There is, nevertheless, a significant body of literature examining the extent to which T is implicated in hierarchy formation in animals; indeed the study of hormone-dominance interactions began with ethological explanations of animal social life. Thus, whilst acknowledging the need to be vigilant about drawing potentially mis-leading parallels, the data can be instructive in attempting to generate new hypotheses about how to think about this relationship in humans, particularly where data exists in non-human primates (Chambers, 1998).

Androgens, and principally T, have important effects upon sexual and aggressive behaviours in many vertebrates (Benton, 1992). Whilst, numerous studies have demonstrated the association between T levels and aggressive behaviour in a wide range of species but these effects are not simple, nor, remotely straightforward. As there are literally thousands of such studies ranging from fish through non-human primates, the interested reader is directed to the following reviews (Bouissou, 1983; Dixon, 1980). The role of androgens in these contexts varies a great deal however, depending upon the species concerned, environmental considerations, and psychosocial factors. With specific reference to non-human primates, studies demonstrate that, in males, T levels are modified in conjunction with a change in status; rising when a dominant position is achieved or defended, and falling when the animal is dominated (Bernstein, Gordon, & Rose, 1983; Eberhart, Keverne, & Meller, 1980). Because these studies concerning the relationship between aggression and T have often explicitly been conducted and considered within a framework of dominance, any simple comparison with humans is complicated by the extent to which findings are variable across animal breeding seasons (Dixon, 1980). Moreover, the literature is replete with contradictory findings; to the extent that it is possible to select studies to buttress almost any position concerning the effects of T and aggression. As Björqvist et al. (1994) remark ‘As far as humans are concerned, the question of whether the level of testosterone in males is related to aggression is more open to debate’ (p.18). Subsequently, whilst there are, no doubt, many facets of the relationship between T and competition in animals that serve to illuminate hormone-competition relationship in humans I wish to focus here on the issue of context.

2.2.1 Context

As with humans, hormones do not completely determine animal behaviour. It has been argued that primates are less dependent on androgens for the expression of aggression than ungulates or other animals lower in the evolutionary chain (Bouissou, 1983). Supporting this position, Wickings and Dixson (1992) and Winslow, Ellingboe, and Miczek (1988) found that administration of exogenous T in squirrel monkeys did not change rank-specific behaviour. However, as mentioned, the relationship is complex and often contradictory. Rejeski et al. (1988) found that intramuscular injection of T-propionate did increase the frequency of aggressive behaviour in monkeys. Monaghan and Glickman (1993) report that under certain conditions, and in a variety of animal models, aggression is dependent on the presence of T and somewhat independent of experience. Under alternate conditions, however, effects of experience often rival hormonal influence (Guhl, 1968). In chickens, for example, other factors such as time of day or the presence of a hen or rival rooster make a difference to the expression of any aggressive behaviour (Allee et al., 1939). Even though the literature is replete with these inconsistent findings it is generally accepted that, among long-lived mammals with highly organised social systems, variation in social context often overrides any hormonal effects on aggression.

Cashdan (1985) reasoned that under circumstances where dominance rank is determined by physical competition, traits such as assertiveness and toughness might be expected to enhance success in competition. This leads to an expectation of a positive correlation between androgen levels and a high dominance rank. However, in stable groups, this is not necessarily the case; often it is the low ranking males who possess the highest levels of T (Sapolsky, 1993). As Cashdan notes, 'Inconsistencies may arise...because behaviours successful in one competitive arena may not be the best route to success in another' (ibid. p.355). Studies involving rodents, for example, illustrate the importance of utilising an appropriate test location when attempting to demonstrate a hormonal response to a range of female behaviours – human or otherwise. It is not possible, for example, to use the criterion rodent resident-intruder paradigm with female rats and mice as subjects, since they do not aggress against each other in that particular test situation (Björkqvist, 2001). Females tested under male conditions, that is conditions in which males demonstrate aggressive behaviour in a natural context, were not aggressive. Only when females were tested under conditions

that mimicked the naturalistic context of aggression, was a relationship demonstrated (Monaghan & Glickman, 1993). As noted earlier, whilst acknowledging that any similarity between behaviours in widely divergent species can be misleading (Mazur, 1973), these findings do at least alert us to the possibility that in human studies involving simple comparisons between males and females, context may be a salient feature requiring consideration.

2.3 Testosterone and Male Competition

It has been suggested that human males, in a variety of different situations, exhibit a characteristically predictable and stereotypical T response, both prior to and following competition. Several studies have found an anticipatory response; with T being found to rise shortly before competition (Booth et al., 1989; Campbell et al., 1988; Mazur, Susman, & Edelbrock, 1997). Following competition, levels of T in the winners either rise or appear high relative to those of losers for a period of time ranging from minutes to days (Booth et al., 1989; Campbell et al., 1988; Elias, 1981; Mazur & Lamb, 1980).

Whilst the majority of the studies conducted so far involve physical exertion, additional studies demonstrate that this dynamic T-behaviour relationship also occurs in non-physical competition. Among the test-situations considered have been contests of reaction time (Gladue, Boechler, & McCaul, 1989) and chess matches (Mazur et al., 1992). Indeed, active participation in the competition may not necessarily be required. For example, T levels have also been shown to increase among spectators watching their favourite sports teams win and drop in fans of the losing teams (Bernhardt et al., 1998). Cohen et al. (1996), in a more naturalistic study, found that T rose in subjects confronted with a symbolic challenge from an insult. Additionally, Mazur and Lamb (1980) found that, following graduation, medical students experienced a rise in T when their mood was elated (although T remained stable if mood did not improve). Thus, in men, the stereotypical T response appears in both non-physical as well as physical competition, in response to symbolic challenges, and status changes generally.

In considering the general nature of these findings it is as well to be aware that a number of factors mediate both the pre-game rise in T and the subsequent hormonal

response to outcome. For example, these apparently characteristic patterns of T response are less likely to occur if mood elevation is absent, if the individual regards the event as unimportant, if the outcome is interpreted as due to luck (Gonzalez-Bono et al., 1999), and if there is a perception that an opponents skill level is particularly different (Mazur, Booth, & Dabbs, 1992; Mazur & Lamb, 1980; Salvador et al., 1987). Collectively, these qualifications suggest that whilst the link between competition and T in males appears to operate in a wide variety of competitive situations, this ‘...is highly contingent on perceptions that gain or loss of status is at stake’ (Bateup et al., 2002, p.183).

Moreover, there are noteworthy exceptions that cast doubt on the claimed stereotypical T response to competition. Five studies found no difference in T levels between winners and losers (Gonzalez-Bono, Salvador, Serrano, & Ricart, 1999; Mazur et al., 1997; Passelergue & Lac, 1999; Salvador et al., 1987; Suay et al., 1999) and one study found losers to have higher T levels than winners (Filaire et al., 2001).

Any simple comparison of the findings amongst these studies is complicated by methodological variance. Whilst there is, as Dabbs (1997) states, ‘a sameness to the research on testosterone’ (p.557), there are fundamental methodological differences which limit useful inter-study analysis and also, importantly, limit interpretations that can be usefully drawn from each individual study. As might be expected, individual authors differ in the degree of emphasis they place on aspects of the relationship. However, the studies themselves follow a very similar pattern. Pre-competition T levels are determined (and occasionally baseline data is also collected), followed by a competitive encounter, and then finally post-competition T levels are established. In the following sections I discuss aspects of these studies in greater detail and subsequently consider various experimental confounds, illustrating their relative importance in contributing to the results obtained. Whilst these points all deal with studies involving males, where the same points apply to the two studies with females these are treated separately under the section on female competition (2.4).

2.3.1 Basal vs. Reciprocal Model

Early attempts at determining the relationship between behaviour and T were typified by a methodological approach termed the basal model. In its most basic guise, this

model argued that basal levels of circulating T - pre-determined by a range of factors such as the genes - were assumed to yield a particular magnitude of behavioural response. In many respects this approach has been superseded by the reciprocal model. Mazur and Booth (1991) have, amongst others, argued strongly that hormone-behaviour relationships should be viewed within the context of this reciprocal model. They contend that, in addition to T leading to dominating behaviour, winning and losing (especially if continued) leads to changes in circulating levels of T. These changes subsequently modify the probability that an individual will behave in a particular fashion on a subsequent occasion.

In non-human subjects, the model used to study the effects of T on behaviour has been to manipulate the hormone level, either by injection of exogenous T or (in males) by removal of the main T producing organs (the testes), and then observing behaviour as the dependant variable (Bernstein et al. 1974; Sapolsky, 1993; Svare, 1983). As Mazur, Booth, and Dabbs (1992) wryly note, however, 'For human subjects this method is neither practical nor safe' (p.70). Consequently, much of the research focusing on the relationship between hormones and behaviour has concentrated on the effect that dominance encounters have on circulating T. It is worth noting that in their response to Mazur and Booths' target article on T and dominance, Christenson and Breedlove (1998) suggested that, whilst the reciprocal model is undeniably correct, a synthesis of the two models is likely to produce a '...fuller explanation of the phenomena' (p.367). They maintain there must be a limit on the extent to which androgen levels are affected by behaviour and that individual differences still do exist in basal levels of T. For example, whilst highly variable there are still gender differences in T levels, with males generally having higher levels than females. It is within the context of this reciprocal model, however, that this thesis examines endocrine responses to competition.

2.3.2 Anticipatory Effects and Hormone Variability

During instances of physiological and psychological stress circulating levels of T typically fall as levels of F rise. However, what characterises the hormone-competition relationship is that anticipatory response to competition includes elevations of F (Filaire et al., 2001; Passelergue & Lac, 1999; Suay et al., 1999) *and* elevations of T (Booth et al., 1989; Mazur, Susman, & Edelbrock, 1997; Suay et al., 1999) which directly

contrasts with the relationship one would ordinarily see in response to a psychological or physiological stressor.

Even though certain authors insist on asserting that, in a variety of physical and non-physical conditions, male T varies in predictable ways before competitive matches (cf. Booth et al., 1989; Campbell et al., 1988; Suay et al., 1999) the literature is, at best, equivocal on this matter. For example, even in a study by Mazur, Susman, and Edelbrock (1997) which is frequently cited as having demonstrated that T rises prior to competition, there is a far from uniform finding. Indeed, as a cautionary note, the authors inform us that 25% of their sample demonstrate a decrease in T prior to competition, reporting ‘...the anticipatory rise, while present in the aggregate, is not highly reliable across subjects’ (p.322). Additional authors have also been unable to replicate a pre-competition rise in T. Filaire et al. (2001) found no significant differences in T determined from a resting state to pre-competition, as did Salvador et al. (2002) when they amended their analysis and replaced individual scores with group mean data. Given the confusing nature of these findings our concern must be to determine the extent to which the contradictions evident in the literature reflect actual effects or are products of methodological disparity. This is a task which, as yet, remains to be undertaken.

If levels of T and F vary in relation to competition then there clearly has to be a level from which they vary. However, the notion of a stable hormone baseline is somewhat erroneous. T levels are under the control of genetic, environmental, and socially mediated factors, and this control seems to ensure that, at least in males, T concentrations are subject to considerable episodic fluctuation and random variability. Fluctuations also occur across seasons, age, possibly race and, importantly, in both sexes there is diurnal variation (Dabbs, 1990). Additionally, in the case of females, it has been argued that T varies across the menstrual cycle. It has been argued that the wide episodic fluctuations in circulating steroid levels may mean that analysis of a single sample is useful only in medical procedures such as screening. In this respect, Hofman (2001) has argued that multiple sampling is a pre-requisite for accurate assessment of bio-behavioural measures. And yet, baselines in hormone-competition studies have tended to be determined, if at all, by the collection of just one sample. Over and above this questionable approach, any attempt to interpret the position that levels of

circulating T show a stereotypical response preceding and following competition is made additionally problematical by methodological variance in sampling times. Salvador et al. (1987) draw attention to the difficulty of interpreting results across studies with respect to diverse temporal collection times, stating that ‘...a more detailed study of the temporal course of these variations [T] is required’ (p.13). Chapter 4 is expressly concerned with the issue of single time-point sampling and consequences of failing to accurately determine baseline data in hormone-competition studies.

2.3.3 Normal Stress Response

Mazur (1985) proposed a biosocial theory of status in face to face interaction; hypothesising a feedback loop between an individual’s T level and his or her assertiveness in attempting to achieve or maintain interpersonal status (the reciprocal relationship discussed earlier). Many of the hormone-competition studies have been conducted within the context of this theory. Within the framework of the reciprocal model Mazur hypothesised ‘...each win reinforces a high T level which reinforces further competitiveness. Conversely, each loss produces a drop in T, which may be functional in inhibiting the loser from engaging in further potentially damaging competition’ (p.70). Modification of levels of circulating hormones (T and F) are thus seen as having a functional significance that fundamentally underpins hierarchical status. One of the central tenets of this model, and a mechanism that has been proposed as operating across a range of competitive situations, is the manipulation of stress levels (Brinkerhoff & Booth, 1981; Mazur & Booth, 1998). Several authors have considered that competition, or behaviour intended to dominate, is in some way implicated with stress. It has been suggested, for example, that due to methodological weaknesses the relationship between T and dominance could be interpreted as a response to stress rather than competition *per se*. And yet, as noted in the previous section, one feature that typifies the endocrine-competition response is that it appears to differ from the response observed in anticipation of other psychologically stressful situations.

In an attempt to explain their findings that T rose following exercise but decreased following competition, Salvador et al. (1987) stated, ‘It is well known that stress (physical or psychological) produces decreases in serum testosterone’ (p.12). Going on to state, ‘Therefore the decreases found in the competitive situation can be interpreted as a reaction to the stress experienced in this competition that outweighs the

increases occasioned by exercise per se' (p.12). However, whilst utilising what initially appears to be an intuitively appealing explanation – stressors such as surgery or military training can indeed lead to reductions in circulating levels of T (Kruezer et al., 1972) – their explanation actually seems at odds with the nature of the relationship being described. It is precisely that T does not follow a traditional model of stress that differentiates this hormone-competition relationship from a standard stress response. Their sampling regimen also suggests uncertainty over the validity of the findings. A look at their methodology reveals the initial measure of T, which presumably was intended as a baseline measure, was taken only 10 minutes before exercise. If the competition was being taken seriously, then the literature in exercise and performance psychology informs us that competitors are, at this stage, likely to be in a state of high anxiety/stress and so this pre-competition measure cannot, with any degree of confidence, be assumed to represent a stable baseline measure against which increases or decreases can be gauged.

2.3.4 Confounding Effects of Exercise

Several studies, unrelated to the investigation of hormone-dominance relationships, have revealed that levels of both androgens and corticosteroids change following exercise (Dessypris, Kuoppasalmi, & Alderkreutz, 1976; Frey, 1982; Kraemer et al., 1998; Lincoln, Rowe, & Racey, 1974; Sutton et al., 1973; Vogel et al., 1971). Despite these numerous reports about the effects physical exercise can exert on T production, Zitzmann and Nieschlag (2001) note that comparative interpretation of their findings is made difficult, as 'Many of these studies lack control groups, have very low numbers of participants, or combine different effectors on hormone levels' (p.183). Elias (1981) suggested, 'Little is known about the mechanisms involved in exercise-induced gonadal and adrenocortical activity' (p.221). Whatever the mechanism involved, it is clear that when considering the dynamic relationship between T and competition, studies involving a physical component are presented with a particular methodological challenge. And yet, as Salvador et al. (1987) remark, '...sporting encounters and fighting encounters involve an important degree of physical exercise, whose influence was not controlled in earlier studies' (p.10).

Six years earlier, Elias (1981) had clearly noted these difficulties and proposed two alternative hypotheses to account for his findings that, whilst T levels were higher

in the winners than the losers of a wrestling match, the losers also experienced elevated levels of T. Firstly, Elias argued the experience of victory may lead to increased levels of T and F which are superimposed on the purely physiological increase of the steroids brought about by physical exertion. And secondly, the experience of defeat may reduce circulating levels of T and F; the decline being counteracted by the large increase brought about by physical activity. The result being a net increase in steroid levels that is lower in losers than among winners. Though Elias neglected to mention it, these effects are likely to be additionally superimposed on changes in circadian activity of unknown magnitude.

In the studies appearing after Elias' work it appears that, in failing to address methodological considerations for the physical component of their studies, no-one sought to follow up on his hypotheses. However, some years later, during a study involving manipulation of outcome in a computer-driven reaction time task, Gladue, Boechler, and McCaul (1989) obliquely addressed Elias' concerns by attempting to answer two critical questions which, at that time, had not been satisfactorily dealt with. Firstly, their study sought to determine whether elevations in T levels occur during competition without a physical component and secondly, in response to Elias suggestion that changes in T and F levels may be the result of wholly separate controlling mechanisms, '...do all hormones respond to dominance-status change; that is do both testosterone and cortisol change in tandem with outcome of competition for status?' (p.411). The first of these questions was crucial because it was, and remains, important to know if subsequent changes in T level are a by-product, '...of physiological arousal and metabolic excitation of endocrine functioning, in which the exercise generally elevates T levels for all competitors, and then 'winning' somehow maintains those levels' (p.411).

In a recently published study, Salvador et al. (2002) make an additional important point, seemingly not considered elsewhere; that the physical effects of warming up have not been considered. This additional feature of the exercise-hormone relationship seeks to further undermine current methodological practices and subsequently the veracity of the claim that hormones (and particularly T) show a stereotypical response to competition. With concerns over how context can profoundly influence the expression of the hormone-behaviour relationship in mind, it is entirely

appropriate to examine these relationships within the framework of physical encounters. Unfortunately, the methodological difficulties encountered have yet to be adequately addressed. Consequently, in order to circumvent some of these complications, the principal experimental study reported in chapter 5 employs a non-physical dyadic encounter.

2.3.5 Outcome and Cognitive Interpretation

Sade (1973) informs us of the comparative ease with which it is possible to observe and to quantify social dominance and fighting behaviour in non-human primate species such as the macaque. Going on to suggest that, although this is due to the limited stereotyped motor patterns used to signal aggression and fighting, more important is the highly ritualised fighting behaviour; a style of interaction which enables winners and losers to be identified unambiguously. However, the concept of victory and defeat in humans may be more problematic than either Booth et al. (1989), who suggest that tennis matches are 'ideal' for this type of research, not least because they produce '...clear winners and losers' (p.558), or Mazur, Booth, and Dabbs (1992) who state '...athletic contest...is a convenient situation for study...and has a clear winner and loser' (p.71) appear to acknowledge.

We can legitimately ask: how useful is it to examine the effects of T in humans on the basis of contest outcome? That is, can victory and defeat be objectively determined given that humans appraise and interpret the events in their lives, so that winning becomes a relative judgment representing a subjective experience. In this regard, Schultheiss et al. (1999) ventured, the discrepancy in findings between studies casts doubt on '...the general efficacy of situational factors such as contest outcome to influence testosterone levels' (p. 234), further suggesting that '...personality factors may moderate individuals testosterone responses to succeeding or failing at a dominance contest. Specifically, the strength of an individuals need for dominance or status may play a crucial role in how the individual responds hormonally to dominance outcomes' (p.234).

Other authors have also considered similar possibilities. Salvador et al. (2002) suggested that complex psychological processes related to emotional and/or cognitive interpretation of the situation may be more important for hormonal responses than the

outcome itself. In their study entitled 'Testosterone and chess competition' Mazur, Booth, and Dabbs (1992) addressed the same point in a novel manner, by investigating the closeness of the win/loss dichotomy. Recently, Schultheiss et al. (1999) found that T showed a positive correlation with motivation to win and Gonzales-Bono et al., (1999) found that T correlated with external attributions of outcome. Thus, if it possible to demonstrate that T is more closely related to interpretation or attribution of events than it is to a conceptually naive notion of victory, then previous work in this area may require re-interpretation.

An ancillary point arising from the literature concerns the unit of analysis. In their study of six male tennis players, Booth et al. (1989) applied the terms winning and losing to the outcome of singles matches (rather than doubles matches which they also played at the same time) because singles matches were argued to 'gain more status for the team' (p.559). Notwithstanding several other methodological limitations in this study, the fact that individual responses were measured on the basis of team status makes any interpretation of these data problematical.

2.3.6 Mood

As an alternate, or perhaps supplementary, explanation for the equivocal nature of findings, a number of studies have investigated the role of affective states as a variable mediating hormonal response to competition. Whilst a number of studies have neglected to consider mood or performance attribution (i.e. Elias, 1981) and specifically for females, the study of rugby players conducted by Bateup et al. (2002), both Mazur and Lamb (1980) and Gladue, Boechler, and McCaul (1989) suggest that mood may play an important, and perhaps critical, mediating role between status achievement and changes in levels of T. Mazur and Lamb (ibid.) state, '...if an individual with rising status felt emotional elation, his testosterone would rise; but if the change in status were not accompanied by a change in mood, there would be little change in testosterone' (p.237). Similarly, Gladue, Boechler, and McCaul (ibid.) further emphasise that '...a status change with no mood change would have no associated endocrine change' (p.411). These suggestions seem plausible; Heritage et al. (1980), for example, found evidence that catecholamine neurons are target sites for androgens in the brain. This is important because catecholamines are implicated in influencing mood states.

However, as a result of their sampling strategy, which was limited by having to rely on the collection of blood samples, Mazur and Lamb's (1980.) finding that T was correlated with mood following graduation actually represented a mood measurement taken twenty-four hrs after the graduation ceremony. In light of both the variability of male T and lack of explanatory mechanism, this finding may reflect an *a priori* desire to unearth a relationship rather than evidence of a valid relationship. In the study conducted by Gladue, Boechler, and McCaul (1989), the authors make explicit reference to the ambiguous nature of their findings from the mood data; in part this results from having collected data on mood following competition only. Thus, differences between winners and losers on scores of depression involved the implicit and questionable assumption that mood states among all competitors prior to competition were identical or extremely similar. Whilst both these findings have been used to support evidence that T rises in conjunction with mood, there are serious questions concerning the validity of the findings.

Whilst the general association between emotion and endocrine function is a close one, there is reason to be cautious when attempting an interpretation of findings that suggest changes in T levels are correlated with mood. Specifically, mood is extraordinarily transient, and given the nature of steroid metabolism described earlier in the chapter, a specific measure of mood is unlikely to reflect a specific and highly accurate measure of T; the latency is simply too uncertain. It is difficult to envisage a way to overcome this methodological difficulty and this theme is discussed further in chapter 5. Finally, and importantly, in suggesting that T levels are responsive to competition only where mood changes are evident severely diminishes the case for a relationship between T and competition, intimating that the relationship could easily exist without recourse to social dominance and competition at all. Further research is required to elucidate this extremely important issue.

2.4 Human Females

A number of possibilities have been considered for expecting dissimilar hormonal responses to competition between female and male competitors. Amongst others, Bateup et al. (2002) proposed several potential reasons. Firstly, they noted that as

women produce five to seven times less T than males it is unlikely to have a potent effect on female behaviour. Secondly, they proposed that resulting from the differing sources of androgen production between men and women it is biologically plausible to believe that women differ from men in their endocrine response to competition.

We presently understand less about the role of T in females than in males, both physiologically and psychologically. The organisational effects of T on the central nervous system are widely thought to become fixed at some pre-adult stage in life, pre-dating subsequent activational effects. Nevertheless, Brain and Susman (1997) have suggested that the influence on neural architecture might not be limited to the developmental stages of life. If correct, this suggests an even more sustainable link between androgens and competitive behaviour. And yet, in an important review article concerning the relationship between T and male dominance, Mazur and Booth (1998) suggested that female T levels are not influenced by competition, either prior to or in relation to successful status outcome. However, their evidence, based on two studies (one of which was unpublished) appeared weak; and, in a subsequent response to the article, Kemper (1998) urged caution, believing that the lack of peer reviewed studies was a problem and the data were insufficient to have any confidence in their judgement.

At that time, the only published experimental study to consider hormonal responses to competition in women had been conducted by Mazur, Susman, and Edelbrock (1997). Their protocol involved collecting salivary samples from both men and women ($n = 28$ and 32 respectively) at baseline, pre, and post-competition phase of a basic video game. They found that in the pre-competition phase, men's T rose, but they were unable to demonstrate a post-game difference in T levels on the basis of outcome and attributed this finding to the lack of differentiation between mood states in winners and losers. They additionally suggested that men may not have found the video game sufficiently engaging to invest in the experience as a challenge to their status. F levels, whilst individually highly variable, fell systematically in the group mean from the first until last sample.

In the female participants the response was different. They appeared not to experience a significant change in levels of either T or F, either pre or post-competition, in spite of winners experiencing an increase in mood. In actual fact T levels fell

systematically from the first until last collection event. Interestingly, female F levels appeared higher than the mens and a number of interpretations were proposed to explain this finding. The authors believed this finding may indicate that women experienced the event as more unfamiliar or challenging than did the men. A second tentative explanation was that the falling trend in F was actually a function of falling T levels. This study represents a constructive addition to the hormone-competition literature, in as much as it was the first published study of the female hormone-competition relationship. The study possessed a relatively sound methodology with a reasonably comprehensive sampling regimen, and identified issues surrounding the contribution of affective state. However, it is worth considering what limitations, if any, might impact upon the interpretation of the results.

One commendable aspect of the methodology was the attempt, through the use of Hemastix[®], to control for possible blood contamination in the salivary samples. *All* samples were found to be contaminated. In order to ascertain whether this blood contamination exerted a significant affect on salivary concentrations of T, the authors performed a correlation between Hemastix[®] and raw sample scores. Unfortunately, this well-meaning procedure is essentially meaningless. Where a high Hemastix[®] score correlates with a high T level there is simply no way of ascertaining if this level would have been high anyway, especially if levels are only marginally higher than they might ordinarily have expected. Given that chronobiological changes in T are precisely what the study is attempting to ascertain, the fact that they reported all the samples were contaminated with blood essentially invalidates the results; especially when salivary T levels were found to fall in concert with falling blood contamination. The study results could, for example, equally be explained away as simply a drop in the amount of blood contamination in the salivary samples. The following chapter, which describes issues surrounding the use of salivary samples in bio-behavioural research programmes, discusses this issue in greater detail. A further, potentially serious, shortcoming was the failure to ascertain whether female participants were administering the contraceptive pill. One of the following sections (2.4) describes in some detail the importance of excluding female participants from hormone-competition studies where they are found to be administering hormonal medication.

More recently, Bateup et al. (2002) conducted a naturalistic study involving physical competition and female rugby players. Over and above determining whether an endocrine response occurred, either pre or post-competition, they explored several issues related to perception. As with Mazur, Susman, and Edelbrock's study, there are a number of methodological concerns which make interpretation of the findings problematic. For example, one potentially troubling aspect of this study concerns salivary collection. In what appears to be an otherwise commendable collection procedure, no mention is made of participants first rinsing their mouths with water prior to salivary collection. As highlighted in the Mazur study, contamination by blood is a problematic aspect of salivary measurement and not one that can simply be ignored. This is particularly true where a study employs an experimental paradigm including a highly physical, high-contact component, such as rugby. It is entirely possible that following a rugby match – which Bateup et al. (2002) describe in their own words as 'aggressive' – that salivary samples had become contaminated by blood. The fact that the authors subsequently went on to re-analyse samples with unusually high levels of T employing 2 and 4 times dilution, rather than considering that these were samples contaminated by blood, gives serious cause for concern about the quality of their hormonal data.

Moreover, whilst reporting a 24% increase in T between baseline and pre-competition, a *t*-test revealed that the comparison between normalised T scores at baseline and pre-competition didn't in fact reach significance ($p=0.06$). The rise of 49% from pre to post-match, whilst reaching significance ($p=0.002$), is far higher than any comparable findings in males. Consequently, it is worth considering further. The post-match salivary sample was reported as being collected immediately following the match, and yet the feedback mechanisms for production of female T, described earlier, preclude any immediate response. It is generally assumed that any modification of androgen levels is likely to take around 30mins from stimulus. Of course, having played an 80 minute game of rugby complicates matters somewhat. Nevertheless, the relationship tested is one concerning T and outcome to competition, and not merely physical exercise *per se*. Moreover, given that the increase occurred irrespective of outcome it seems to suggest that at least some of this increase is a result of the physical exertion involved in the game. For example, Kraemer et al. (1998) found a 38%

increase in T following resistance exercise – albeit in a male population. Thus, an initially impressive finding appears questionable.

2.4.1 Exclusion Criteria

In clinical studies examining themes such as hormone replacement therapy and menopausal status, several researchers have expressed regret that occasionally, in selecting participants, a greater emphasis is not placed upon exclusion criteria (Braunstein, 2002). In hormone-competition studies, exclusion criteria have also been afforded a lower priority than would seem necessary to ensure an accurate assessment of hormone-behaviour interactions, especially where those studies involve female participants (i.e. Mazur, Susman, & Edelbrock, 1997).

Over and above changes brought about by psycho-social or environmental factors, levels of the free fraction of T are particularly susceptible to fluctuations in the levels of the binding globulin (SHBG) in the blood. SHBG levels are dependant upon a range of physiological and pathological factors in humans and are increased by conditions such as hyperthyroidism, the contraceptive pill (particularly where these are high oestrogen preparations), cirrhosis of the liver, and age. Levels may be reduced with hypothyroidism, glucocorticoid applications, and obesity (Braunstein, 2002; Demers, 1995). Presumably with knowledge of at least some of these factors, several authors have endeavoured to consider participant selection for hormone-competition studies based on factors overtly affecting T levels. However, whilst attempting to control for possible menstrual variance by obtaining self report data about menstrual cycle phase, Mazur, Susman, and Edelbrock (1997), in the one experimental paper examining female endocrine response to competition, neglected to enquire whether female subjects were administering the birth control pill (neither, for that matter, did Bateup et al., in the only other published study on hormones and female competition). Although, in order to justify this omission, they cite a study by Kirschbaum and Hellhammer (1994). This paper demonstrates that the birth control pill exerts ‘...no discernible effect on salivary C’ (p.321). However, the thrust of Mazur et al.’s article explicitly concerns T. Consequently, their comment, whilst no doubt accurate, is somewhat disingenuous. Administering the various forms of birth control pill has been found to impact critically upon salivary T levels in females. This occurs because levels of SHBG, to which T is strongly bound in the blood, can be profoundly altered. Moreover, as Guay (2002)

notes, 'birth control pills will not only increase SHBG but decrease ovary's production of hormones, especially androgens' (p.S84). Thus, studies where participants are administering the pill, and which have attempted to determine the free fraction of T, are made that much more difficult to interpret. In order to circumvent these difficulties particular attention is paid to exclusion criteria during participant selection in this thesis.

2.4.2 Menstrual Variability

The human menstrual cycle involves a complex and regular change in female anatomy and physiology, and several studies have reported that T varies in a predictable manner throughout the course of a menstrual cycle; although the magnitude and subsequent relationship to behaviour is, at present, unclear. In a study examining plasma T changes across the menstrual cycle, Vermeulen and Verdonk (1976) stated, 'It is evident from this study that T...plasma levels do show statistically significant cyclical variations with maximal variations around ovulation' (p.493). Just over ten years later this finding was echoed in a study by Morris, et al. (1987), in which they indicated a rise in T levels around the mid- point of the cycle. Dabbs and de La Rue (1991), whilst finding a mid-cycle rise in salivary T, suggested that, as this variation was smaller than circadian variation, 'Menstrual cycle effects can be ignored in most research relating psychological and behavioural variables to individual differences in testosterone' (p.182). The issue of menstrual variability, particularly in the free fraction of T, still requires elucidation and our laboratory is currently analysing data from such a study.

2.5 Data Analysis

In describing the pattern of results in hormone-competition studies there are few, if any, instances where inferential procedures are carried out on raw data. Therefore, in considering the findings from the literature it is not inconsequential to consider how the data have been treated and how this might impact upon interpretation of study findings. Hormonal data generally and T specifically are typified by often extreme inter-individual variation, with high standard deviations tending to obscure any trends or effects within the group mean. This is particularly true given the necessarily limited number of participants in naturalistic studies of competition. Even where results have been used to support evidence of a hormonal response to competition these findings

often do not actually reach significance (i.e. Mazur & Lamb, 1980; Booth et al., 1989). Researchers have subsequently adopted a range of approaches in the analysis of their hormonal data. One way in which researchers have attempted to overcome the challenges posed by inter-individual variation is by not only looking for significant differences in group mean data or describing their findings subjectively but by additionally examining percentage change from the mean. (i.e. Elias, 1981; Salvador et al., 1987). Moreover, the manner in which data sets are analysed exerts a powerful effect on study findings and hence interpretation. For example, whilst Elias (ibid.) found a significant change between T levels measured 10 minutes prior to a match and levels measured 10 minutes after the match, this was a percentage change. The effect disappeared when raw scores were analysed as group mean data: 'Because of the considerable individual variation in baseline [T] levels, mean values...were not significantly different...However, *percent changes* in [T]...were significantly different' (Elias, ibid. p.217)

In an attempt to overcome these difficulties and render the samples suitable for meaningful analysis several authors have normalised their data. Mazur and Lamb (1980) treated the first post-match (study 1) serum sample as an arbitrary baseline, subtracting this T level from all remaining scores for any individual. Given that they had collected a pre-competition sample it is unclear why this wasn't the sample treated as baseline and, unfortunately, they don't indicate why. Whatever the reason, it has the effect of ensuring that all pre- until immediately post-match levels of T fall when in their raw state this seems extremely unlikely. Adapting this method slightly, Mazur et al. (1992, 1997) and Bateup et al. (2002) divided raw hormone scores by an individual's highest recorded level. These normalised values produce a range from just above zero to a maximum of 1. However, whilst Mazur et al. (ibid.) claim that raw data produce similar results they neglect to report them. Salvador et al. (1987) employed an alternate method. They used percentage change in order to '...eliminate the masking effect of large individual differences' (p.10). Other authors have applied a logarithmic transformation to their data in order to correct for skewness. All these approaches are fundamentally valid. However, the difficulty comes when the data is manipulated to such an extent that results no longer bear any resemblance to the raw data.

Where claimed baselines have been established with one sample (see chapter 4 for a discussion of difficulties associated with this approach) a number of authors (i.e. McCaul, Gladue, & Joppa, 1992) have appropriately employed ANCOVA analyses with baseline values acting as the co-variate. However, where ANOVA procedures have been employed to analyse data, authors have committed themselves to reporting all data points. This often proves difficult in behavioural-endocrine research as samples frequently fail to meet inclusion criteria, i.e. blood contamination, lack of sample, or unpredictable behaviour of the sample during storage or assay; this is particularly true of earlier studies who had little relevant literature to guide their collection and storage methods. Unfortunately, whilst SPSS will employ default data to fill missing data cells, the manner in which it accomplishes this is not always the most appropriate (Keppel & Wickens, 2004). Presented with this obstacle Mazur, Susman, and Edelbrock (1997) attempted to overcome the problem of missing data by calculating the mean of the corresponding values from all other subjects and substituting that value. For example, if 3 subjects collected samples at 1, 2, and 3pm and in one subject the 2pm sample were missing, the mean value of the other two subjects at 2pm were utilised. Whilst it is easy to see why they have chosen this step, particularly in instances where only one isolated sample is missing, it is actually extremely unsatisfactory. They employed this procedure to insert the missing data from four female samples; the missing data were replaced by other subjects who fell into the same winning or losing condition. Given the explicit reference to extreme inter-individual variation it seems inappropriate to utilise scores for other subjects. This procedure effectively serves no purpose other than to strengthen the existing relationship. A much more satisfactory approach would have been to replace missing scores with a neutral estimate from the specific individuals whose scores are absent. For example, by obtaining the mean value of the 2 scores either side of the missing variable, and utilising that. Alternatively of course, where more than one data point is missing, remove participants from the analysis entirely.

2.6 Cortisol and Competition

Compared with T, information about the role of F in competition is not as comprehensive, in either males or females (Bateup et al., 2002). Studies suggest, however, that F may play a role in a constellation of behaviours generally believed to be

important in (physical) competition; including aggression, arousal and mobilisation of physiological resources to deal with an impending threat or challenge. In line with these suggestions, there is evidence of anticipatory elevation in male F levels prior to competition and a rise during the competitive event itself (Booth et al., 1989; Elias, 1981; Filaire et al., 2001). Despite these correlational findings, F response may be related to something other than competition. In the study conducted by Filaire et al. (ibid.) increases in F occurred in both winners and losers, with no significant difference between the two during any collection event. They suggested, ‘...that C [cortisol] is related to “situational” stress and arousal rather than status changes’ (p.60). Another aspect of the link between F and competition is that top-seeded male tennis players exhibited consistently lower F levels than did less talented players, suggesting that highly successful competitors may have above average ability at coping, or managing stress (Booth et al., 1989). It was for this reason that the experimental procedure described in chapter 5 utilised a game in which skill could not reasonably be expected to play a part.

2.7 Summary

The relationship between social dominance and androgens in males, although not particularly straightforward, is seemingly well established in a wide range of non-human species, including primates. Several studies have utilised sporting competitions in order to examine similar endocrine changes in humans. Several of these studies appear to demonstrate a similar pattern of response, with T levels rising if an individual is victorious and falling if defeated. Yet, there are additional studies that are unable to replicate these findings. Moreover, there are methodological flaws in these human studies of sufficient magnitude to give a great deal of pause before the claim is substantiated with any degree of confidence. With T associated very much with males there is little comparable research on human females. Of the two published studies conducted with human females, one finds evidence of an endocrine response and the other not. These two studies invariably suffer the same methodological limitations as their male counterparts. Given that T levels in females are in the order of eight-ten times lower than in males and given that measurement of free T is often inaccurate, these methodological flaws get thrown into sharp relief, reducing even further our confidence

in the findings. It should also be noted that, as a result of the often subtle nature of female social dominance hierarchies, the link between hormone-competition studies and social dominance in females is even less certain than it is in males

This review chapter has illustrated that any attempt to extrapolate findings from male studies onto females may be conceptually naïve and likely to lead to an erroneous understanding of the true nature of the relationship between T and social dominance. Issues such as context are critical in unearthing when a hormone-dominance relationship might exist and when that relationship will be subsumed under greater environmental or psycho-social pressures. The following chapter takes one of the issues identified here as problematic - the measurement of T in women - and explores the theme in greater detail. Chapter 4 subsequently addresses another theme identified here – sampling strategies. Chapter 5 is an experimental chapter that uses methodology incorporating improvements in methodology to many of the limitations addressed in the present chapter.

Quantification of Salivary Testosterone: Optimising and Validating an In-house Enzyme Linked Immuno-Sorbant Assay

3. Introduction

Quantitative determination of circulating testosterone (T) is possible from an assortment of biological material; i.e. plasma, serum, hair, saliva, and urine. Principally though, there are two fluids which have been widely utilised for research purposes in the behavioural sciences; plasma and saliva. Early attempts at determining concentrations of circulating T, primarily for clinical purposes, traditionally utilised plasma or serum. Whilst this was advantageous in facilitating direct analysis of the circulating concentrations of total T in the blood, for the purposes of examining chronobiological changes there were, and remain, several distinct limitations. As blood samples require time-consuming and often stressful venipuncture, obtaining invasive multiple samples over a period of hours can be painful and this procedure can be unattractive to participants (Dabbs, 1990). Moreover, measurement tends to be of the total rather than free, biologically active, fraction of T. The measurement of this free fraction is essential in bio-behavioural research as it represents, in part, the fraction of total T that is available to exert a biological action at the cellular level (Smith et al., 1979).

3.1 Free vs. Un-bound Testosterone

Throughout this thesis the distinction is drawn between free, bio-available, and total T. Although the figures reported vary somewhat (partially as a reflection of changes in assay technology), in women it is widely accepted that approximately 60-70% of T circulating in the blood is bound strongly to Sex Hormone Binding Globulin (SHBG); women have a higher percentage of T bound to SHBG than do men (Guay, 2002). An additional 30-40% is weakly bound to the small protein Albumin (Braunstein, 2002; Carlstrom & Rannevik, 1990). As Albumin is a small protein (molecular weight of

approximately 69,000) it can pass through the salivary membrane. Conversely, SHBG is a large carrier protein (molecular weight of 100-200,000) and so cannot easily pass the salivary membrane. In an early paper, Vermuelen and Verdonk (1972) proposed that the free fraction of T may be between 4 and 10%. However, although it may well prove to be inaccurate, it is now generally accepted that in females this figure is likely to be around 1-3% of circulating T free to exert its biological action (Demers, 1995; Vermeulen, 1998), with a greater amount being bio-available. That is, under certain conditions the bio-available T dissociates from its carrier protein and becomes free. Besch et al. (1982) state, 'Currently, it is generally held that only the non-protein-bound or 'free' fraction of a hormone, such as a steroid enters the cell of the target tissue and interacts with its specific receptor protein' (p.221). As an adjunct to this latter point it should be noted that around 50% of pre-menopausal (and almost 100% of post-menopausal) T is synthesised by peripheral conversion, distinct from any endocrine function (Labrie et al., 2000).

Even now, simple methods for determining free steroid concentrations in plasma have not been fully developed. Consequently, most current procedures involve technically demanding and time consuming centrifugal ultrafiltration or equilibrium dialysis (Carlstrom & Rannevik, 1990; Hammond et al., 1980; Riad-Fahmy et al., 1982); processes which in themselves yield results with varying degrees of accuracy. In an effort to circumvent the problem of not easily being able to measure free-T some authors use mathematical equations, such as the free androgen index, mass action formulation, or the Soodegaard equation. All are mathematical formula of varying complexity, for working out the free component of T based on the measurement of total T and, as a minimum, SHBG. However, because these equations rely on an affinity constant (SHBG in the case of T) and because estimations of the affinity constant varies widely, there are very real problems with these approximations (Besch et al., 1982). As Vermuelen, Verdonck, and Kaufman (1999) point out, 'As the total serum testosterone concentration is subject to variations in the concentration of the binding proteins, it is not a reliable index of bioavailable testosterone' (p.3666).

Faced with these challenges, there has been a growing awareness of the potential value of utilising saliva for measuring hormone concentrations (James & Baxendale, 1982; Mandel, 1993). Consequently, the use of saliva as a diagnostic tool in bio-

behavioural and clinical research has grown significantly during the previous two decades (Quissell, 1993). Salivary sampling regimens have several obvious and distinct advantages over blood sampling. They accommodate frequent, easy, collection by non-invasive, relatively stress free-techniques, thereby facilitating ‘...short term dynamic tests, pharmacokinetic analyses, and studies of chronobiological changes’ (Riad-Fahmy et al., 1982, p.367). In addition, it has been reported that the majority of subjects find little difficulty in salivating directly into collection containers, providing adequate volumes (between 3mL and 5mL) for determining a steroid hormone profile in less than 5 minutes (Dabbs, 1991).

3.2 Blood-Saliva Correlation

As steroid-binding proteins cannot easily pass through the salivary membrane one of the advantages in using saliva as the biological fluid for study centres around the claim that it contains concentrations of analyte similar to, or even identical with, the non-protein bound (free) concentrations in blood (Baxendale, Reed, & James, 1980; Kirschbaum & Hellhammer, 1994; Smith et al., 1979; Vittek et al., 1985; Walker et al., 1980; Wang, et al., 1981). These claims generally relate to males, however. Although there are fewer studies examining this association in females, it has been suggested that correlations between T levels in serum and saliva may be significantly lower in females than it is in males (Granger et al., 1999). Shirtcliff, Granger, and Likos (2002) recently published a study which appears to cast doubt on the veracity of this correlational relationship in females, stating ‘Regardless of assay method, salivary testosterone levels are modestly correlated with serum levels for males but not necessarily females’ (p.1798). The absolute concentration, whilst reflecting accurately the unbound fraction in the plasma (and also the fraction which is not bound to SHBG) is approximately twice the concentration of free-T in plasma, in contrast to the findings in male subjects (Baxendale & James, 1982; Smith et al., 1979). One major implication of these findings is that substitution of saliva for serum T levels in bio-behavioural studies may estimate the T-behaviour relationship differently for females than males.

3.3 Normal Testosterone Ranges

In women there is still little accurate evidence about what constitutes a normal range for T (Davis & Tran, 2001; Sina-Hikim et al. 1998). A selection of studies are reported in Table 1 illustrating that, within hormone-competition studies, reported values for T

show a wide variation. It may be that at least some of this variation results simply from different subject pools. However, it is plausible that much more of the variation comes from differing assay protocols.

Table 1 Example of Ranges for Male and Female Testosterone in Hormone-Competition Studies

Author	Sex	T Levels	Medium	Assay
Salvador et al. (2003)	♂	128.56-311.2 pmol/L	Saliva	RIA
Bateup et al. (2002)	♀	Not reported	Saliva	RIA
Mazur et al. (1997)	♀	72.81-90.14 pmol/L	Saliva	RIA
Elias (1981)	♂	3.61-33.12 nmol/mL	Plasma	RIA
Gladue et al. (1989)	♂	191-955 pmol/L	Saliva	ELISA

Given the extensive range of factors affecting assay performance, any simple comparison of T levels across different assays is problematic. For example, because of the high sensitivity and specificity of the ‘in-house’ assay optimised for use in this thesis, measured values are lower than the direct salivary measurements with commercial kits. In a recent symposium entitled *Women and Androgens* at the annual scientific meeting of the Endocrine Society, Barrett-Connor (2004) acknowledged the difficulties with simple inter-laboratory comparisons, going on to suggest that where levels reported from a laboratory are within accepted limits, and that there are no contra-indications, then each lab should set its own normal range. Our laboratory has followed this advice, and mean T levels are reported in the studies of chapters 4 and 5.

3.4 Contamination of Salivary Samples

Clinical researchers have established that collection techniques can affect the integrity of salivary samples, which subsequently interferes with ability to accurately determine hormones levels. For example, given the difference in T concentration between blood and salivary samples, leaking of blood or serum into the mouth (i.e. due to gum disease, injury, caffeine intake, or vigorous cleaning) can affect the integrity of quantitative estimates of salivary concentrations (Granger, 2000, 2004). Whilst the contamination of salivary samples by blood is widely recognized as a potential problem in the scientific fields driving assay development there appears to be something of a malaise regarding the issue in hormone-competition studies. In response to a question about the incidence of haemoglobin contamination in salivary samples, at a round table discussion on assay development and collection procedures (Tenovus workshop, 1982), Schürmeyer

suggested that researchers might reasonably expect a contamination incidence of around 5-10%. With literally a couple of notable exceptions, however, most authors have neglected to report how they have dealt with this serious confound.

In contrast to the majority of other authors, Mazur et al. (1980, 1987) and Gladue, Boechler, and McCaul (1989) paid careful attention to the issue of contamination, adopting what, at the time, must have appeared a sound approach. In utilising measurement strips called Hemastix[®] they attempted to ascertain which, if any, of their salivary samples may have become contaminated with blood. Unfortunately, an important and timely paper, just published by Kivlighan et al. (2004), casts severe doubt on the suitability of this approach. They state, '...quantitative measurements of testosterone in saliva are sensitive to the effects of blood contamination. Importantly, the confounding effects of blood leakage cannot be adequately screened or controlled by visual inspection of sample discoloration or using the Hemastix[®] approach' (p.45).

Whilst new methods have, in the past year, been developed to assess blood contamination in salivary samples (i.e. Salimetrics now offer an assay to determine blood contamination in salivary samples), *all* samples in this thesis were assayed for F in addition to T, even when not reported. At the time the present studies were conceived, and with no evidence to the contrary, a supposition was made that contamination of salivary samples by blood would show up very obviously in F results. For example, where a typical level for female salivary F in our lab may be in the range of 0.9-15 ng/mL, a contaminated sample would produce levels at over 30ng/mL. In light of the findings by Granger et al. (2004) it is difficult to maintain confidence that the manipulation check employed actually served its function effectively. Several samples with F level in excess of 50ng/mL were detected. These were clearly contaminated and subjects were subsequently removed entirely from corresponding analysis. However, as early attempts at collecting and analysing samples were not unproblematic the protocols were gradually enhanced to the extent that blood contamination might be considered as effectively reduced to an incidence of below 1%. For example, following suggestions by scientists such as Adlercreutz (1990), our subjects were asked to refrain from brushing their teeth prior to sample collection; more recent guidelines from companies such as Salimetrics (2003) suggest that subjects should not brush teeth for upto two hours prior to collection. Subjects were additionally asked to rinse their mouths



thoroughly prior to collection, not to eat, not to smoke or drink caffeine. These steps effectively reduced the amount of samples seen in the assays with abnormal levels. Therefore, I am confident that whilst this manipulation check may not allow detection of marginally contaminated samples, our stringent collection procedure significantly reduces the contaminated samples encountered.

In chapter 2 I proposed that single time-point sampling acts as a potentially fatal methodological confound in a substantial number of the hormone-competition studies to date. Whilst this issue is explored in greater detail in chapter 4 it is worth noting here that potential contamination of salivary samples by blood further complicates the use of single samples. In chapter 2 I additionally noted that where samples have been found with aberrant levels, researchers have treated them by diluting the samples rather than consider contamination as an issue (i.e. Gladue et al., 1989). Whilst this step may be appropriate for male salivary samples, dilutions of female salivary samples which may already contain levels of T at the utmost sensitivity of the assay are rendered undeterminable by dilution.

3.5 Collection Methods

With the increased use of salivary measures in bio-behavioural and clinical research, several research teams have sought to advance our understanding of the circumstances and conditions that may influence the validity of these assessments (e.g. Lipson & Ellison, 1989; Shirtcliff et al., 2001). And yet, as Shirtcliff, Granger, and Likos (2002) remind us, ‘...there remain wide gaps in information available to guide researchers in the utility of this potentially important method’ (p.62).

Typically, saliva is collected by having a participant deposit between 3-5mL into a collection container; this step is usually reported as taking between 3 and 8 mins. The saliva is then stored frozen prior to assay. While clinical subjects may be willing and able to provide un-stimulated samples that can be immediately frozen, this protocol is often unpractical for field collection (Lipson & Ellison, 1988). As such, collection of saliva samples under these conditions may necessitate certain changes from common clinical practices. In respect of this, studies have shown the actual method of stimulating saliva production can potentially distort T levels in a range of assays. Yet, with this in

mind, authors of hormone-competition studies have not always been as stringent as they might in reporting conditions of collection.

Although collection of saliva is often referred to as un-stimulated, this is somewhat misleading. If the salivary gland is un-stimulated it does not produce saliva; what is meant essentially is saliva produced by minimal stimulation (Read, 1989). Small amounts of saliva can be collected without the need to externally stimulate production, but the 3-5mL usually collected often requires subjects to stimulate saliva production in some way. This issue is particularly pertinent when investigating changes in concentration of salivary hormones prior to competition, where time constraints exist and feelings of stress or anxiety may make saliva production more troublesome. Where necessary, salivary flow can be stimulated by a variety of agents. The use of a range of stimuli has been reported; i.e. paraffin wax, rubber bands, sugar-free gum, cotton swabs, and citric acid (Malamud, 1993; Navazesh, 1993).

Table 2 Methods of Salivary Stimulation in Hormone-Competition Studies

<i>Study</i>	<i>Sex</i>	<i>Collection Method</i>
Salvador et al (2003)	♂	No stimulation
Bateup et al. (2002)	♂♀	Sugar-free Original flavour Trident Gum
Filaire et al. (2001)	♂	No stimulation
Passelargue (1999)	♂	Not specified
Gonzalez Bono et al. (1999)	♂	Water and Lemon
Schultheiss et al. (1999)	♂	Sugar-free gum
Mazur et al. (1997)	♂♀	Sugar-free gum*
Mazur et al. (1992)	♂	Crystals of Sodium Chloride
McCaul et al. (1992)	♂	Citric acid drops
Dabbs, (1990)	♂	Sugar-free Original flavour Trident Gum

* The authors report that *nearly* all subjects chewed gum to stimulate saliva

Importantly, Vining et al. (1983) described that non-polar analytes are released into saliva through membranes via a mechanisms that is not flow dependent. Thus, unlike salivary IgA, T concentrations in saliva have been reported as independent of flow rate (Riad-Fahmy et al., 1982). In bio-behavioural studies investigating hormones and competition a range of techniques for stimulating saliva have been employed, and these are indicated in Table 2. Perhaps just as troubling as those authors that have utilised techniques which previous research has shown to be problematic (in the sense that they distort T levels) are those who have not reported methods of collection at all.

3.6 Assay Techniques

Quantitative determination of circulating T can be derived utilising a variety of different procedures such as; radioimmunoassay, luminescence immunoassay, equilibrium dialysis, gel filtration, protein precipitation, gas-chromatography mass-spectrometry, centrifugal ultrafiltration, and enzyme-linked immunosorbant assay (Edwards, 1985; Kemeny, 1991; Sinha-Hikim et al., 1998), each with their own inherent strengths and weaknesses. However, once a medium in which to analyse T has been selected (saliva in the case of this thesis) the question then becomes what type of assay should be employed. Although the bewildering array of techniques can make selecting the most appropriate test perplexing, Edwards (*ibid.*) tells us, 'A systematic and objective approach will...indicate the most appropriate technique for the particular application in mind' (p.2).

In both saliva and blood, RIA was until fairly recently the method of choice for determining concentrations of circulating T. Indeed, despite a number of limitations that will be discussed presently, traditional analogue RIA is still widely used (Rosner, 2001). Several authors, however, have ardently expressed concern that this method is insensitive at the lower end of the T range; precisely the range in which female salivary T falls. In this respect, Sinha-Hikim et al. (1998) state, '...the existing assays, designed for the measurement of serum testosterone levels in men, lack the sensitivity required for the precise measurement of the low levels prevalent in women' (p.1312). Indeed, in a variety of research arenas, accurate measurement of female free-T is widely regarded as problematic. In a clinical paper on screening for androgen insufficiency Guay (2002) noted, 'A major problem in assessing T [female] is the inaccuracy of the measurements by current assays. These assays were developed for men who have higher levels of circulating T' (p.S83). Recent articles published by Taieb et al. (2002) and Herold and Fitzgerald (2004) illustrate that when measured by automated processes, T concentrations may be inaccurate to the order of magnitude of 500% [*sic.*]. This issue - one of measurement - therefore acts as one potentially serious limiting factor in the confidence that we can place in the results of hormone-behaviour studies in females. Although as technology advances and measurement continues to become more straightforward there is still a need for highly sensitive, reliable, and efficient immunoassays with accessible reagents and materials for the determination of T in females (*c.f* Granger et al., 1999; Guay, 2002).

Given the difficulties associated with accurately measuring female free-T, the enzyme-linked immunosorbant assay (ELISA) protocol has a number of features to recommend it over RIA. For example, reagents are cheap in comparison with RIA, the laboratory in which I undertake my research employs staff with substantial expertise in ELISA development, there is no hazard of radiation, and because of the low concentrations of antigens in saliva, HIV and hepatitis infections are much less of a danger from saliva than from blood (Major et al. 1991). Indeed, unless visibly contaminated with blood, human saliva is not considered a class 2 biohazard, affording researchers and institutions administrative and safety benefits (U.S. Centers for Disease Control). As a consequence of these factors the ELISA method stands out as the most appropriate method for use during work conducted in relation to this thesis.

There are a number of features that are required of an effective assay. As Kemeny and Chantler (1988) note, 'The type of assay should be closely tailored to the particular task for which it is required' (p.2); and the requirements of diagnostic laboratories are often very different than those of bio-behavioural research labs. For the purposes of this thesis (i.e. assessing chronobiological changes in the very low levels of free salivary testosterone in females) issues such as ease of use and speed are less important characteristics than sensitivity and accuracy. Subsequently, the aim of this chapter is to describe the optimisation and validation of an extremely sensitive in-house ELISA, designed specifically for determining salivary free-T in women. In particular, the ELISA is evaluated for its accuracy, specificity, and precision. The following sections describe the fundamental principals behind the ELISA principle and several aspects of assay parameters optimised: temperature of incubation, incubation times, enzyme dilution, antibody quantity, and specificity.

3.1 ELISA Principle

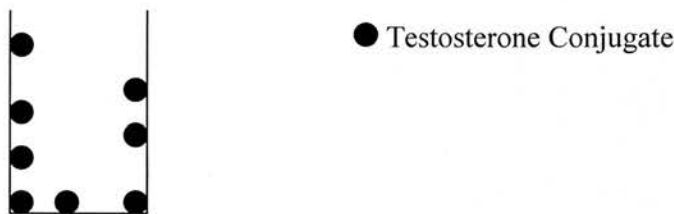
The principle of the assay optimised in this thesis is based on the in-direct, competitive binding technique. O'Sullivan, Bridges, and Marks describe enzyme assays as '...a group of protein binding assays in which the molecular recognition properties of antibodies are used' (1979, p.221). Essentially, the T present in salivary samples competes with a fixed and limited amount of T coated on the micro-titre plate, for

binding sites on an antibody. Because the concentration of the T coated to the wells is held constant, while the concentration of T in the salivary samples vary, the amount of enzyme labelled second antibody bound to the first antibody is inversely proportional to the concentration of the unlabelled analyte present in the sample. In order to facilitate a clearer understanding of the process the following section provides a description of the basic ELISA procedure utilised in this thesis.

3.1.1 Procedure (adopted in our laboratory)

Step One: Coating the Plate

Proteins spontaneously bind, strongly and non-specifically, to chemically clean plastic or glass surfaces. The majority of ELISA assays depend upon the ability of protein antigen, or antibody, to bind in this manner to the bottoms of plastic wells in 96-well plates. The antigen (in this case Testosterone Conjugate – which is bound to Bovine Serum Albumin) is added to the plate where it attaches itself to the bottom and sides of the well. The plate is covered with parafilm to retard evaporation and incubated overnight at 4°C.

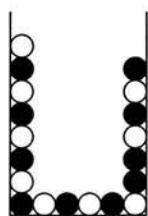


Step Two: Washing

The antigen, which is absorbed onto the well surface, stays in the well whilst any excess antigen not bound to the well walls is removed with wash buffer.

Step Three: Blocking the Plate

After coating the plate (binding a layer of protein), additional layers of antibody or antigen are bound. These, however, must bind specifically only to the protein in the first layer. Therefore, it is crucial to block the plate to prevent spontaneous non-specific binding of these additional layers to the plastic well. One method of blocking utilizes a protein such as bovine serum albumin (BSA) to coat any plastic left exposed after the initial coating. Hence, blocking buffer is added to the wells, filling any gaps on the well walls and base. This step reduces non-specific binding.



● Testosterone Conjugate
○ Bovine Serum Albumin

Step Four: Blotting

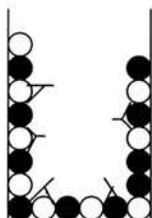
Following incubation for 1 hr. at 37°C any excess blocking buffer is discarded and the plate blotted dry.

Step Five: Adding Samples and Standards

T standards and saliva samples are added to the wells. Because the zero standards contain no T, all the antibody (see following step) binds to the conjugate coated onto the well; hence the high colour upon completion of the assay procedure.

Step Six: Adding Antibody

Antibody (Sheep Anti-T) is added to the wells and subsequently attaches only to the antigen. All the antibodies in the sample must be attached to an antigen to get an accurate measurement. At this point the antibodies are invisible and as in preceding steps any excess antibody is washed out leaving only the antibodies attached to the antigens. At this stage a biological sandwich is formed.



Step Seven: Washing

Following incubation at 37°C for 1 hr. the plate is washed four times

Step Eight: Adding Enzyme

The enzyme, horse radish peroxidase, which is attached to the Donkey Anti-Sheep antibody, is added to the wells. The conjugate contains an enzyme which indicates the presence of the antibody.

Step Nine: Washing

Following incubation at 37°C for 1 hr. any excess conjugate is washed out leaving the enzyme markers attached to the antibodies.

Step Ten: Adding Substrate

Substrate is then added. This part of the process is a chemical reaction. The substrate reacts with the enzyme on the conjugate and produces a blue colour of varying intensity which can then be measured.

Step Eleven: Stopping the Reaction

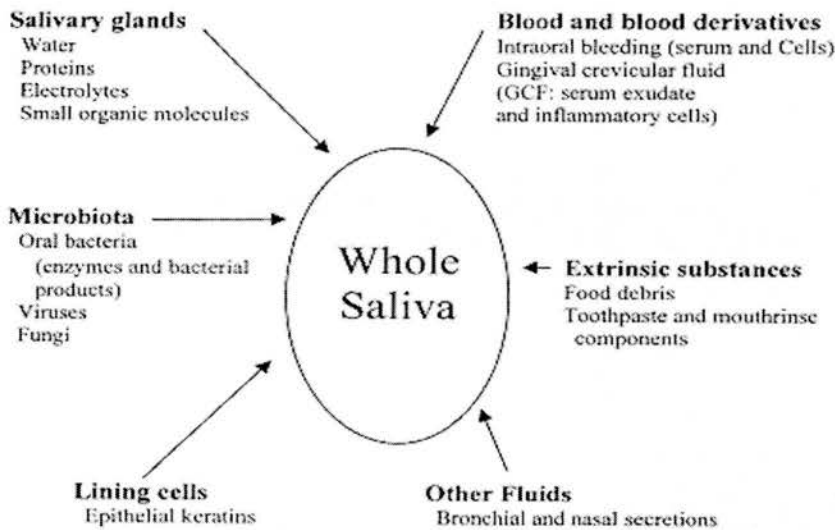
The reaction is finally 'stopped' with sulphuric acid that retards the activity of the enzyme. The timing of this step is critical. If you leave one reaction longer than another, the colour may be more intense because the reaction has had more time to proceed. The process is useful only if each step and concentration of reactants is standardised. The intensity of the colour (optical density, OD) is then measured in an ELISA plate reader and the T concentration calculated from the standard curves produced in each plate.

3.1.2 Direct vs. In-direct Technique

Solid phase assays can adopt either a direct or an in-direct approach to determining levels of T. Whilst several assays, including those commercially available kits, employ a direct method (that is, salivary samples are not treated in any way prior to assay) there are a number of limitations with this approach when examining female androgens. Whilst the benefits of utilising saliva over serum have previously been described, saliva is a far from inert substance (see fig. 2); it contains various contaminants, such as bacteria, leukocytes, mucins, and extremely importantly for enzyme assays, endogenous enzymes. All of which can interfere with assays based on the ELISA technique. As a consequence, salivary samples are rendered extremely susceptible to interfering agents such as pH imbalance which yields results that are, unpredictably, either too high or low (alkaline samples will, for example, tend to yield low results). Schwartz et al. (1998) report that when pH falls below 4 or rises above 9, then assay performance is likely to be compromised. At a round table discussion of sex hormones and corticosteroid assays, Aldercreutz, citing a range of studies, noted '...testosterone assays do not work well in non-extracted plasma' (1990, p.387). Moreover, Jones, Murphy, and Alaghband-Zadeh (2004) reported that '...some samples from female subjects give falsely high results

when measured with direct immunoassay’ (p.51). In order to circumvent some of these difficulties, the method utilized throughout the thesis is the indirect method. That is, prior to assay samples are treated with an extraction step that removes interfering agents.

Figure 2 Components of Whole Saliva. (from Kaufman and Lamster, 2002).



3.1.3 Solvent Extraction

Sample purification, by means of ether extraction, is an extremely well-established assay procedure. Although time consuming it is particularly accurate and has advantages over both heat and acid treatment. Despite both acid and heat denaturing enzymes, rendering them inactive and therefore unable to interfere with the assay (Hubl, 1984), salivary ELISA is extremely sensitive to pH. As it can be difficult to regulate pH following addition of acid it is used increasingly less as a means of purifying samples. Consequently, the extraction step reported in the section on assay validation is one involving di-ethyl ether.

Glass-stoppered quick-fit Fischer glass tubes and ordinary Pyrex glass tubes (10x75mL) were prepared, first by soaking overnight with water and hospital grade detergent and then being put through an intense (hot) dishwasher wash-rinse cycle. In order to remove any remaining proteins that may have bound to the inside of the tubes they were then further rinsed with de-ionised water. Finally, following immersion in

methanol, they were drip-dried. Aliquots of saliva were thawed and re-centrifuged at 6000rpm for 2 mins. in order to break down any remaining muco-polysaccharides. 500 μ l of sample was pipetted into a quick fit tube along with 4mL of ether. The tubes were then vortex mixed for ten minutes and placed at -80°C until the aqueous phase was frozen (usually between two and five minutes). The remaining unfrozen ether was then decanted into ordinary glass tubes. These tubes were placed into a water bath pre-heated to 45°C, where the ether was evaporated (leaving T attached to the tube walls). 500 μ l of assay buffer was added to reconstitute the T. The solution was left to stand at room temperature for 30 minutes to equilibrate and finally vortex mixed for 5mins. prior to use in the assays.

3.2 Introduction

Kemeny and Chantler (1998) note that 'Despite the apparent simplicity of ELISA, the quality of assays developed by different laboratories varies tremendously and this is largely due to the level of assay optimisation undertaken' (p.17). Indeed Kemeny (1991) had previously noted, the most important factor in determining performance of an assay was the extent to which it had been optimised. The following studies represent the attempt to satisfy the requirements of our laboratory for this assay. Because of the extremely low levels of T in female saliva, one of the particular requirements for determining female T is that the assay be especially sensitive. However, the ELISA process sits within a complex web of inter-locked parameters, and achieving this sensitivity requires a constant balancing act between reagents to arrive at and maintain the final protocol. For example, achieving sensitivity comes at the cost of making the assay more susceptible to interfering factors than routine serum methods. One of the ways to minimise the impact of interfering factors is to treat the samples prior to assay and this is described in the previous section. In line with guidelines taken from Al-Dujaili et al. (1988), several aspects of assay optimisation are reported here: temperature of incubation, incubation times, enzyme dilution, volume of the sample, antibody quantity, and specificity.

3.3 Amount of Conjugate

The effects of altering the amount of conjugate in the assay are diverse. Too much conjugate will reduce sensitivity, whilst too little increases the risk of losing the signal. Consequently, the first experiment was designed to investigate the most effective amount of T conjugate capable of producing the optimal optical density (OD) signal. In the case of our assay, this optimal signal is defined as one that produces an OD of 1 following between 10-15 minutes of incubation following addition of substrate.

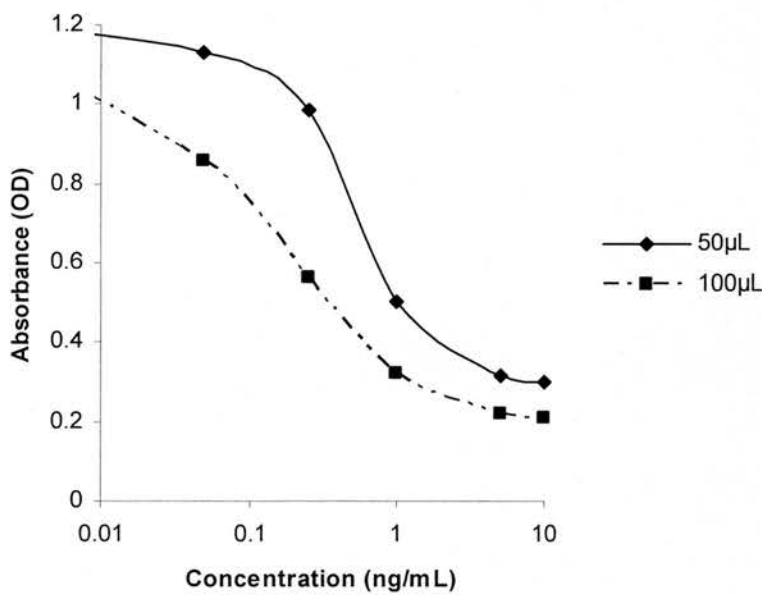
3.3.1 Method

Lane A on a 96-well plate was coated with 100µl/10mL of T conjugate. Row B was similarly coated with 50µl/10mL of T conjugate. The assay was subsequently performed using procedure described in section 3.8.

3.3.2 Results and Discussion

The standard curve data obtained is presented in fig. 3. As the standard curve created using 100µl/10mL of T conjugate produced a steeper curve between the zero standard the 0.25ng/mL standard it was concluded that this was sufficient to produce a satisfactory standard curve that would allow for accurate determination of low levels of T.

Figure 3 Standard Curves Produced by Varying the Amount of Conjugate



3.4 Amount of Antibody

3.4.1 Method

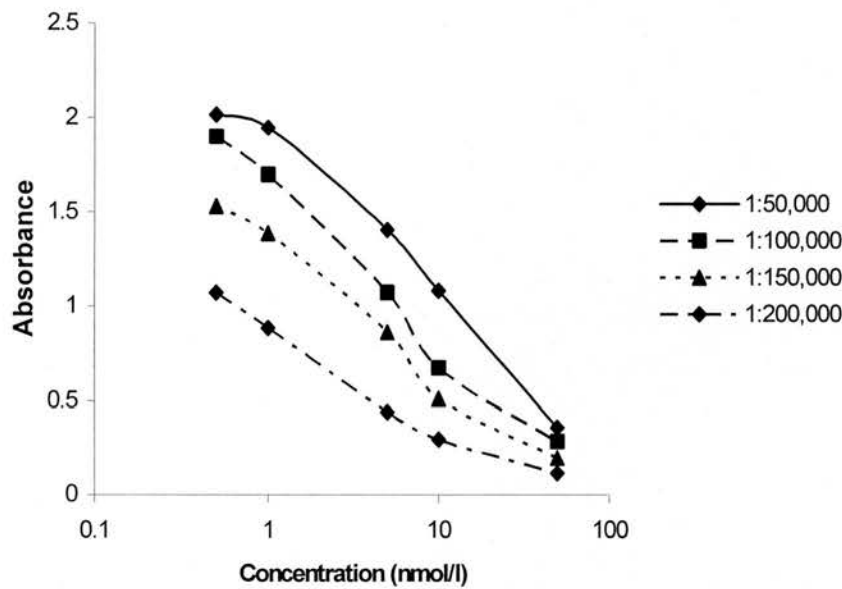
Altering antibody concentration also influences several aspects of assay performance; including sensitivity, signal detection, and precision. In order to examine the most appropriate antibody concentration, lanes A to D on a 96-well plate were coated with

50µl/10mL of conjugate, which represents the optimal concentration as determined during the previous experiment. Other than varying the antibody concentration from 1:50,000 to 1:200,000, the assay was performed using the procedure described in section 3.8

3.4.2 Results and Discussion

Fig. 4 illustrates the standard curve data obtained. It was concluded that antibody concentration of 1:200,000 produces the most sensitive curve with sufficient signal.

Figure 4 Effect of Antibody Concentrations on Testosterone Standard Curve



3.5 Incubation Temperature

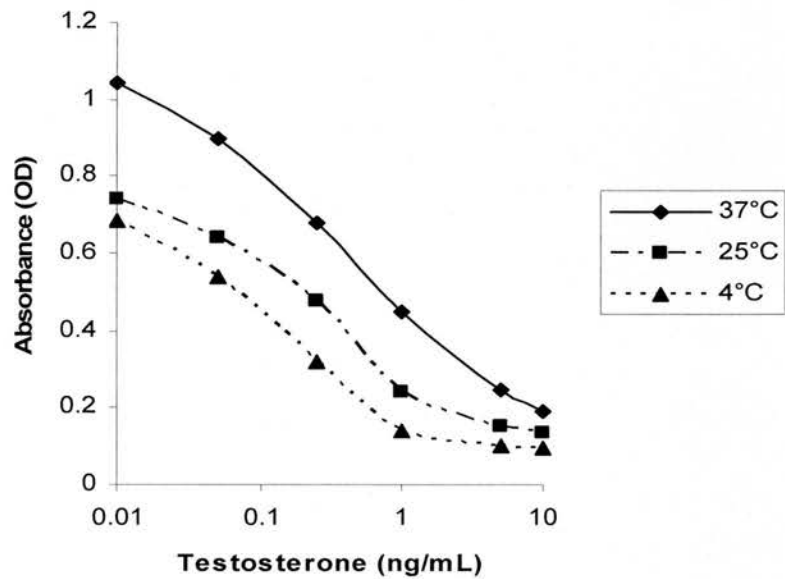
3.5.1 Method

Enzymatic activity is extremely sensitive to heat and the assay is built around enzyme activity. Given that the antibody (Ab) will respond differently at varying temperatures, the assay was incubated at 4°C, 25°C, and 37°C. Three rows of the T standard with an Ab concentration of 1:200,000 were incubated for 1hr. and the procedure was completed as described in section 3.8.

3.5.2 Results and Discussion

Fig. 5 illustrates the highest signal was obtained at 37°C for 1 hr. Therefore, in the optimised protocol 37°C was utilised.

Figure 5 Standard Curves Following Incubation at Varying Temperatures



3.6 Incubation Time

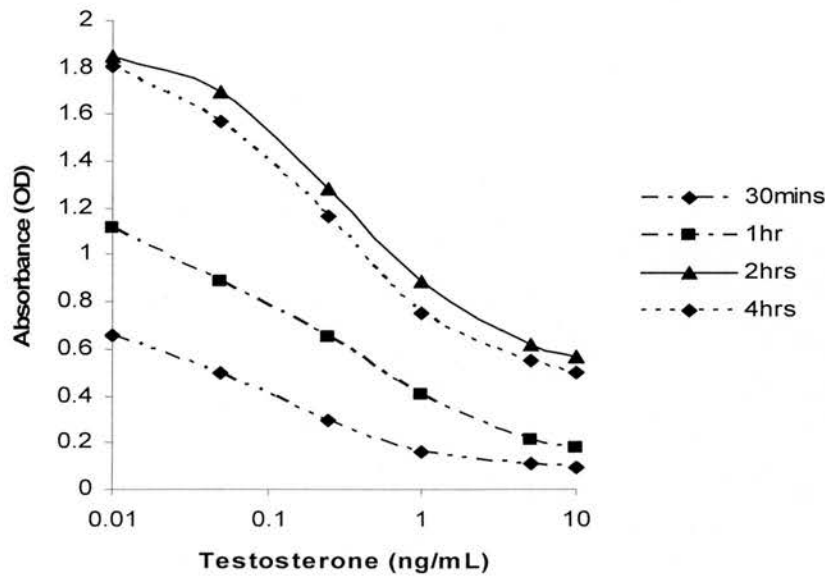
3.6.1 Method

Four lanes of a plate were coated with 50µl/10mL of T conjugate in buffer and antibody concentration of 1:200,000 (as determined in the previous experiment). The plate was then incubated as follows. Row 1 for 30 mins, row 2 for 1hr., row 3 for 2hrs., and row 4 for 4hrs. at 37°C.

3.6.2 Results and Discussion

Fig. 6 illustrates that 1hr. at 37 degrees was sufficient to produce the signal required with a satisfactory standard curve (satisfactory means a steeper standard curve). At 4hrs there was no more binding than at 2hrs. Therefore, incubation for 1hr was sufficient to produce a satisfactory standard curve. Where assays employ low affinity Ab, incubation time may have to be increased upto 16 hrs.

Figure 6 Standard Curves Produced at Differing Incubation Times



3.7 Volume of Sample

Altering the volume of sample assayed impacts upon the accurate measurement of the hormone, particularly for samples containing low levels. The greater the volume of sample the easier it is to measure against the standard curve. Where levels of T are low, as in female saliva, samples are occasionally extracted and concentrated by manipulating re-constitution volume. By contrast, several authors in hormone-competition studies have chosen to dilute samples following high readings that are difficult to interpret.

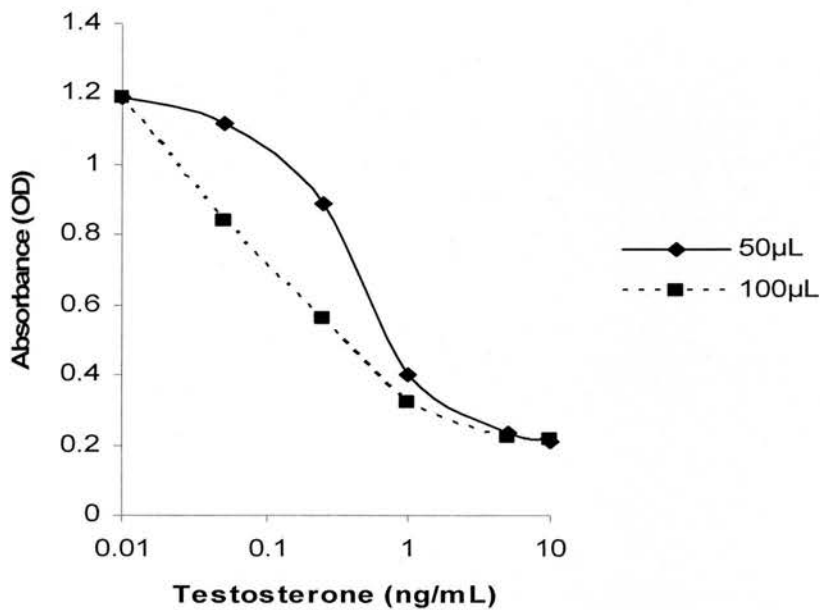
3.7.1 Method

When ascertaining the optimal sample volume, the first two rows (constituting of 24 wells) were prepared with 50 μ l of standard and 100 μ l Ab. In row A standards were pipetted and row B six samples in duplicate were assayed. In rows C and D 100 μ l of the sample and 100 μ l of the antibody were used.

3.7.2 Results and Discussion

Because of low levels of T evident in female salivary samples 100µl was selected. This curve was more sensitive, allowing the measurement of low T levels without recourse to dilution or concentration.

Figure 7 Testosterone ELISA Standard Curves of Varying Volumes of Sample



3.8 Optimised Testosterone Assay Protocol

The final assay procedure has been optimised as follows:

3.8.1 Plate Preparation

- Step 1. Coat plates with T conjugate and leave to incubate overnight at 4°C
- Step 2. Wash 3 times with wash buffer
- Step 3. Block for one hour at 37°C
- Step 4. Discard blocking buffer

3.8.2 ELISA Procedure

- Step 1. Pipette 100µl of standard or previously extracted and re-constituted sample into wells (duplicate). Standards are run at zero, 0.05, 0.25, 1, 5, 10ng/mL
- Step 2. Add 100µl of antibody at 1:200,000 concentration to each well
- Step 3. Shake and incubate at 37 degrees for 1hr.
- Step 4. Discard and wash 4 times (turn plate around each wash)
- Step 5. Add 100µl of enzyme (horseradish peroxidase)
- Step 6. Shake and incubate at 37 degree for 1 hr.
- Step 7. Discard and wash 4 times
- Step 8. Add 100µl of substrate solution (tetramethylbenzidine)
- Step 9. Incubate and shake for 12 mins. at room temperature (cover from light)
- Step 10. Add 50µl of stop solution (H₂SO₄)
- Step 11. Read optical density at 450 absorbance on MRX microplate reader

3.9 Introduction

The validation of salivary steroid assays essentially depends on two independent factors. Firstly, can the steroid concentration in saliva be measured with sufficient accuracy. Secondly, are the levels reported meaningful; that is, do they correlate with the plasma total or plasma free levels, with a criterion method, or with some other physiological parameter of interest (Read, 1989). In order to assess the validity of the in-house ELISA utilised during this thesis the following validation protocols were undertaken: recovery, parallelism, imprecision, sensitivity, specificity, and correlation.

3.10 Accuracy (recovery experiments)

Analytical accuracy can be defined as closeness to a real value, and is affected by a number of factors such as interference from substances other than the analyte, or by steroid loss through extraction and assay. In performing an extraction step as employed with this assay, there is a commensurate loss of analyte. Although there are several criteria against which to judge the efficiency of an extraction procedure, Chard and Martin (1975) suggest the most important of these are 'recovery, reproducibility, and the absence of non-specific effects' (p.75). Therefore, in order to determine how much analyte (i.e. T) is recovered during this procedure the following recovery experiment was conducted.

3.10.1 Method

Salivary samples ($n=20$) collected from female subjects (containing different levels of endogenous T). These samples were extracted using the previously described procedure and then split into 6 groups with 6 samples in each group. Each group was spiked with known quantities of T: group 1: was neat (no spike), group 2: 0.1ng/mL, group 3: 0.2ng/mL, group 4: 0.4ng/mL, group 5: 0.8ng/mL, group 6: 1ng/mL, and assayed utilising the optimised assay protocol. In order to ascertain the final percentage of

analyte recovered the final concentration from each sample was subtracted from the basal value (mean of those samples not spiked with any T) and then divided by the theoretical value (the value obtained from the various spiked concentrations) and then finally multiplied by 100 to give the percent recovery value.

3.10.2 Results and Discussion

The average recovery was 104.0% (range 97.5% to 110%). The mean percentage of T recovered is presented in Table 3. There is very little analyte loss through extraction; the procedure can thus be viewed as successful. Compare these results with from Dabbs (1990) who reported mean recovery of only 80%.

Table 3 Recovery of Known Amounts of Unlabelled Testosterone Added to Salivary Pool

Lane	Added Testosterone Concentration (ng/mL)	Amount Recovered (ng/mL)	% Recovery (Mean)
A	Standard Curve		
B	0.0		
C	0.1	0.095	95
D	0.2	0.21	105
E	0.4	0.39	97.5
F	0.8	0.868	110
G	1.0	1.01	101

3.11 Imprecision

As Challand, Goldie, and Landon (1974) note, ‘An essential requirement of any assay...is that it should be precise’ (p.40). And yet, measurement difficulties resulting from the relatively small molecular size of T and the fact that it is present in female saliva in relatively low concentrations means there is often marked inter and intra-assay bias and variation; this is especially evident in earlier assays but, as a problem, still persist in certain current assay kits. Consequently, comparing T levels determined with different assays and/or different laboratories becomes questionable (Zitmann & Nieschlag, 2001). Direct measurement of free-T concentrations, which are found at much lower concentrations than the total amount of this steroid (around 2%), is subject to even greater difficulties. In addition to error introduced through the optimisation of assay parameters, Ekins (1974) informs us that ‘Counting errors are frequently of less importance than experimental errors in restricting assay precision and sensitivity’ (p.5).

In this context, experimental errors are those incurred through mechanical processes such as washing, pipetting, etc. Two aspects that can affect the imprecision are the timed addition of the substrate and the washing step. In our early work we utilised a single repeating pipette. However, with the purchase of new equipment this was replaced with multiple repeating pipettes. These contributed to improving precision.

3.11.1 Method

Two pools were prepared from salivary samples collected by male and female participants in a variety of experiments. These pools were aliquoted into eppendorf tubes and, in order to determine inter-assay coefficients, one sample from both males and females were included in all assay runs. Intra-assay precision was determined from the mean of 42 replicates assayed in duplicate for male and female T levels.

3.11.2 Results and Discussion

The average intra-assay coefficient of variation was 2.1%. Inter-assay precision of 6.7% was determined from the mean of averaged duplicates for 50 separate runs for male and female aliquots. Both the inter and intra assay coefficients are at levels comparable with the best commercially available assay kits and are significantly better than many assays employed in hormone-competition studies (i.e. an inter-assay CV of 17.1%; Booth et al., 1989).

3.12 Assay Sensitivity

As previously described, without technically challenging steps RIA cannot accurately measure free T at the levels generally found in healthy young women, let alone at levels found post-menopause. As all the work in this thesis is concerned with the measurement of T in women, the assay is required to be especially sensitive. Assays tend to be least precise where they are most sensitive, however, and the reproducibility of replicates powerfully influences lower detections limits. In detecting small physiological changes in hormone levels, Chard and Martin (1974) state that 'In general the lower limit of detection should be half or less of the basal resting concentrations of the hormones' (p.76). The theoretical detection limit of the assay (sensitivity) can be defined as the minimum concentration of T that can inhibit the signal such that it can be distinguished

from the zero standard. Caution is urged, however, when considering sensitivity alone as a marker of assay performance. For example, as Laughlin and Barrett-Connor (2001) note, 'Although the sensitivity of the Sinha-Hikim total testosterone assay is very high, total testosterone levels were apparently measured directly in the dialysate. Equilibrium dialysis would not eliminate cross reacting substances, which may account for higher...testosterone levels with this assay' (p.1844).

3.12.1 Method

Each lane in the micro-titre plate was filled with a known amount of T at the following concentrations: 0.005, 0.0025, 0.001, and 0.0005ng/mL. The plate was then assayed following the optimised protocol.

3.12.2 Results and Discussion

An un-related *t*-test revealed the differences between the zero standard and 0.0005ng/mL concentration were significantly different: $t(1, 11) = 9.098, p < 0.001$. By this method the limit of the assays sensitivity is <0.5pg/mL. Compare this with the sensitivity reported by Filaire et al. (2001) of 15 pg/mL.

3.13 Specificity Studies

Specificity refers to the ability to discriminate between closely related but antigenically different molecules. All assays are, to a greater or lesser extent, subject to interference from other, similar, androgens. Thus, assay results can mislead if what is actually being measured is a T metabolite or precursor, rather than T itself. For example, Turkes et al. (1980) used an assay in which interference by 5 α DHT was negligible, whilst Tames and Swift (1983) used an antiserum having around 60% cross-reactivity with 5 α DHT. As the levels reported were very similar between assays, it is possible that compounds other than 5 α DHT cause interference in immunoassays for T in female saliva. Consequently, a wide range of hormones, similar in structure to T, were tested against the T antibody in order to assess cross-reactivity.

3.13.1 Method

The specificity of the present ELISA was determined by preparing doses from a range of steroids with similar molecular make-up to T and which have previously been shown to interfere with T assays. Each compound was assigned a lane, so 12 determinations were made. They were assayed following the optimised procedure.

3.13.2 Results and Discussion

Table 4 reveals that all the cross-reaction percentages were <25%, indicating that there is no significant interaction of other hormones to the T antibody used in this experiment, although not all known steroid hormones could be tested (Al-Dujaili, 2004).

Table 4 Cross-reactivity between Testosterone and related steroid hormones

<i>Steroid hormone</i>	<i>% Cross- reactivity</i>
Testosterone	100
Nandrolone	22
Dihydrotestosterone	23
Androstenedione*	7.6
DHEA	1.05
DHEA-sulphate*	0.72
Cortisol	0.001
Cortisone	0.03
Pregnenolone	0.05
Progesterone	0.02
Corticosterone	0.24
Prednisolone	0.3
Estradiol-17B	0.52
11-Deoxy-Cortisol	0.2
17-OH-Progesterone	0.02
11-deoxy-Corticosterone	0.06
Cholesterol	0.04

** major precursors for testosterone in females.*

One of the reasons commercial assays invariably produce T results that appear high compared with in-house protocols is because of the cross-reactivity with DHEA-S. The importance of the cross-reactivity findings depends not only on the % cross-reactivity but on the relative concentration of the compound compared against T. For example, DHEA-S occurs in plasma at approximately 100 times the amount of T. Hence, a cross-reactivity of 0.72% is potentially of more importance than the 23% cross-reactivity found in DHT, which occurs at levels below T in plasma. Using the extraction

procedure removed the ability of DHEA-S to interfere with the 'in-house' assay optimised here.

3.14 Comparison between Commercial Salivary Testosterone Kit and 'in-house' Assay

In a series of early studies, Smith et al. (1979) demonstrated the relationship between concentrations of steroids in female saliva and the free, non-protein bound, fraction in plasma. Resulting partly from a lack of expertise in setting up in-house assays, many researchers investigating the relationship between androgens and aggression/dominance use commercially available serum testosterone kits modified for use with saliva (Granger et al., 1999). However, as previously described, these assay kits can impede accurate determination of female salivary T. Nevertheless, in order to determine the comparable performance of the in-house assay, two separate Salimetrics kits were purchased and tested against the in-house ELISA.

3.14.1 Method

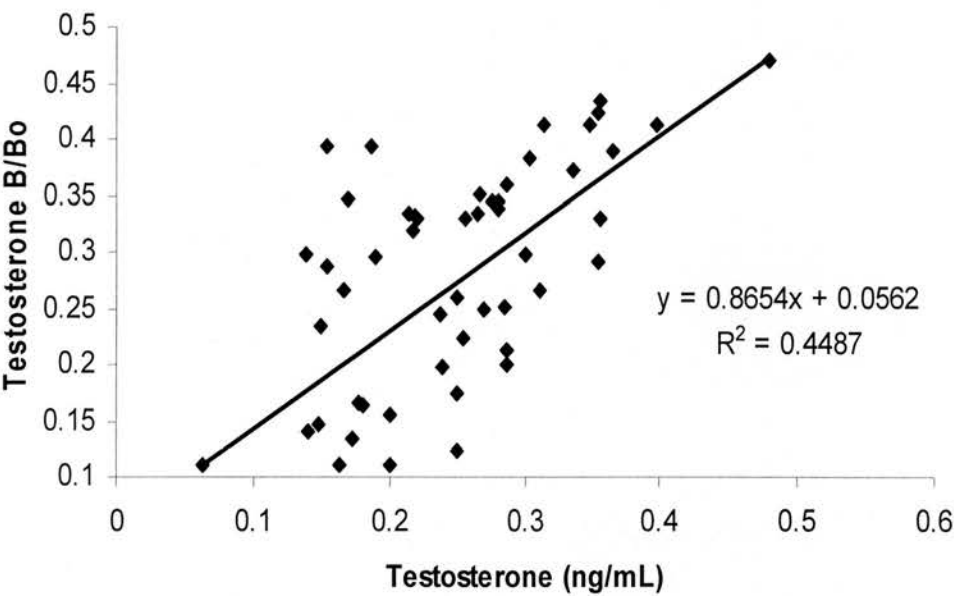
Matched aliquots of saliva were thawed and centrifuged at 6000rpm for 2 mins. Those to be analysed with the in-house assay were then extracted using the step described in section 3.1.3. Following extraction, samples were assayed utilising the optimised assay protocol previously described. Those samples to be analysed by Salimetrics kit were assayed following instructions included with the Salimetrics assay kit. Quality control (QC) samples for the in-house ELISA were prepared from aliquots taken from a mixed female salivary pool which had been previously collected to provide the lab with multiple QC's.

3.14.2 Results and Discussion

It was extremely encouraging to find that levels between the Salimetrics kits and the in-house assay were highly comparable – particularly when calculating the Salimetrics results using their preferred B/Bo equation. Indeed, the agreement between our quality control (QC) samples was excellent at $r=0.97$ (mean in-house assay QC was 0.292ng/mL and the Salimetrics kits QC was 0.285ng/mL). However, the overall level of agreement between the in-house assay and the Salimetrics kits was poor. Previously,

in chapter 3, I drew attention to some of the difficulties faced when using a direct method for ELISA; in particular the problem of pH balance. Salimetrics incorporate a chemical indicator into their assay which detects pH balance in salivary samples; if a sample is alkaline, for example, then it will turn a purplish colour and if acid the sample turns yellow. Salimetrics suggest that samples with a $\text{pH} \leq 4$ or ≥ 9 may be artificially inflated or lowered. Unfortunately, by the time it has turned purple the assay is well under way and there is nothing you can do to change the results. In order to circumvent this difficulty, Salimetrics recommend that all samples are tested for pH balance prior to assay. Of course, the procedure for then remedying pH in-balance then becomes potentially more problematic than the extraction step employed with the in-house ELISA.

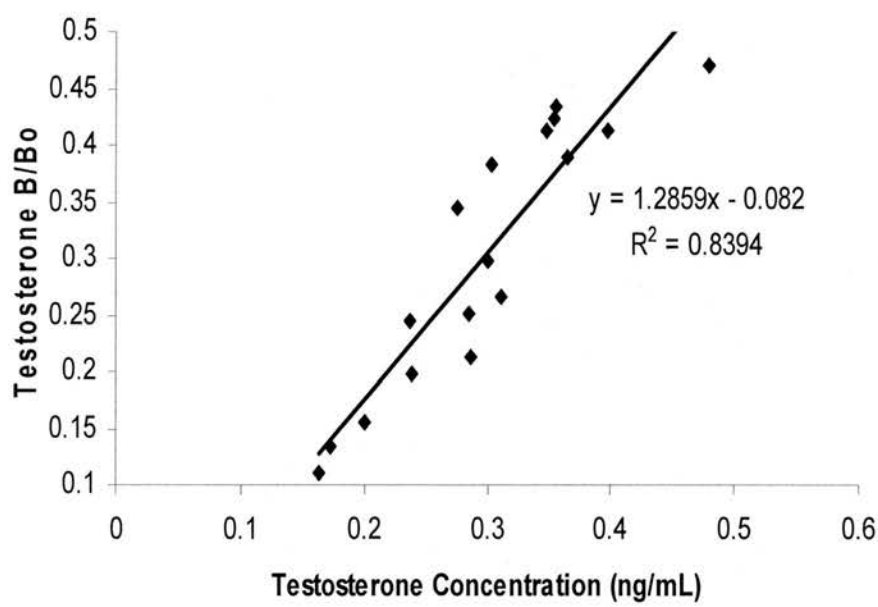
Figure 8 Correlation Between In-house ELISA and Salimetrics Testosterone Assay with Male and Female Salivary Samples



If only a few samples were affected, researchers might consider the comparative simplicity of a commercial kit worth the speed-accuracy trade-off. However, the vast majority of our female samples were revealed to be alkaline and subsequently the results obtained with the Salimetrics kit were questionable. It is this difficulty that we believe contributed to Salimetrics not being able to distinguish temporal changes in

female samples. Fig. 8 illustrates the correlation between the Salimetrics and in-house assays. By contrast, male salivary samples rarely appeared to be alkaline, and this is reflected in a much higher correlation between the two assay procedures (fig.9). Given that it seems likely there will be a higher incidence of alkalinity in female salivary samples, the implications for hormone-competition studies which utilise direct assays is potentially great. The findings from this study suggest, at the very least, that authors place a much greater emphasis on reporting their assay characteristics, and particularly, identify the alkalinity of their samples and what steps they undertook to ensure the assay was not compromised.

Figure 9 Correlation Between In-house ELISA and Salimetrics Testosterone Assay (non-alkaline male and female samples)



3.15 Summary

The chapter reported the development and optimisation of an ‘in-house’ enzyme linked immuno-sorbant assay for the determination of female salivary T. Assay performance was satisfactory, being highly sensitive, accurate, specific, and with good precision. Even though the in-house assay optimised during the course of this thesis is both reliable, accurate and valid, it is indirect. Whilst in many respects this represents a particular strength of the assay, in as much as it removes interfering factors from the

analysis, it is extremely time consuming, labour intensive, and requires particular attention to detail. Our lab currently runs hormonal analysis for several research teams and consequently, time becomes an important consideration. In conjunction with continued assay development there are several new pieces of equipment that are awaiting validation. A plate washer has the potential to further reduce the time spent running the assay whilst improving even further precision. A new plate reader, with improved software, will also provide improved accuracy; thus, development is still ongoing to optimise the protocol.

Circadian Activity, Episodic Fluctuation, and Testosterone Baselines

4. Introduction

In both serum and saliva, testosterone (T) production in healthy young males tends to be characterised by a circadian rhythm, with concentrations highest early in the day and falling away in the late afternoon and early evening (Bremner, Vitiello, & Prinz, 1983; Dabbs, 1990; Goldzieher et al., 1976; James & Baxendale, 1982). In addition to this circadian profile, and resulting from the pulsatile release of gonadatropic releasing hormone (GnRH) from the hypothalamus and luteinizing hormone (LH) from the pituitary, T levels fluctuate, often quite dramatically, in short periods of time (i.e. around 90 minutes). This modulation could, if neglected, act as a potentially serious experimental confound in the design of hormone-competition studies.

In studying hormone-competition interactions several authors neglect to incorporate a baseline into the design of their studies (i.e. Elias, 1981), evaluating instead, the relative difference between pre and post-competition hormone levels. These are often determined by collection of a single salivary (or occasionally in early studies, blood) sample. Where researchers have taken care to account for modulating levels of T in the design of their studies they have tended to control for the circadian component of variation only, by collecting samples at approximately the same time of day and subsequently disregarding episodic fluctuation. Thus, bio-behavioural studies exploring the relationship between T and competition in males have tended to utilise single, one-off, measurements in order to determine T levels at baseline and pre-competition. For example, if a pre-competition sample is required at 11am, a single baseline measurement is also taken at, or near, to 11am on an earlier day. Any deviation from the baseline is thus assumed to be, in some manner, a function of the competitive event or the participant's interpretation of the event. It should be acknowledged that in direct

contrast to the sparse collection of baseline and pre-competition samples, several studies have collected multiple post-competition samples

4.1 Baselines

Establishing a baseline for T is not entirely straightforward. Whilst T production is under genetic control (Meikle et al., 1988) it is also responsive to a range of biological, environmental, and psychosocial stimuli, the relative impact of which are poorly understood. Although not an exhaustive list, amongst those factors identified as having a role in modifying T levels are: fasting, dietary component, sexual activity, alcohol, competition, behaviour intended to increase status, aggression, physical exercise, cognition, stress, immune function, and mood. In line with this, T concentrations have been shown to vary with time of day (Ahokoski et al., 1998; Baxendale, Reed, & James, 1980; Dabbs, 1990; Turkes et al., 1980; Walker et al., 1980), season of the year (Dabbs, 1991), sexual activity (Morris et al., 1987; Swift, 1984; Zitmann & Nieschlag, 2001), and they fluctuate, in males at least, in a pulsatile fashion over minutes and hours (Veldhuis et al., 1987). Moreover, among women, T concentrations are thought to increase around the middle of the menstrual cycle (Massafra et al., 1998; Vermeulen & Verdonck, 1976). Dabbs (1990) encapsulates some of the difficulties in designing suitable sampling regimens for bio-behavioural studies when he states, 'All this variability introduces error into behavioral studies, where stable measures are needed to characterise individual differences and changes over time. Without more information on these changes one cannot know how many participants to run, how many measurements to take, and when to take measurements' (p.83).

Establishing a meaningful T baseline depends, at least in part, on what purpose the baseline is required to serve. In certain clinical practices, for example, researchers have advocated collecting single samples each day over the course of a week and pooling the samples to provide a weekly average. In order to ascertain if an individual has levels of T that might indicate a risk of some clinical condition this approach is entirely reasonable. However, when attempting to examine chronobiological changes in relation to competition this sampling protocol would be unsuitable. In an attempt to establish an appropriate baseline measure researchers interested in hormonal responses to competition have, with few exceptions, utilised only one approach – to collect one-off salivary samples. Within that approach a variety of sampling strategies have been

utilized. Table 5 indicates the sampling regimens used to determine baseline and pre-competition levels of T.

Table 5 Testosterone Baselines: Sampling Times in Hormone-Competition Studies

Author	Sex	Baseline	Pre-Competition
Bateup et al. (2002)	♀	24hrs prior	15 mins prior
Mazur et al. (1997)	♂♀	10 mins prior	3 mins prior
Filaire et al. (2001)	♂	3 weeks prior	5 mins
Gonzalez-Bono et al. (1999)	♂	None established*	45 mins prior
Passelergue and Lac (1999)	♂	15 mins	immediately
Suay et al. (1999)	♂	45 mins prior	immediately
Eubank et al. (1997)	♂	None established*	10 mins prior
McCaul et al. (1992)	♂	None established*	immediately
Booth et al. (1989)	♂	Approx. 24hrs prior	15 mins prior
Gladue et al. (1989)	♂	10 mins prior	5 mins prior
Salvador et al. (1987)	♂	None established*	10 mins prior
Elias (1981)	♂	None established*	10 mins prior

*note: where these authors have not collected baseline data they have effectively classified their pre-competition sample as a baseline

As is evident, authors have not adopted a unified approach to sampling times. And yet, fairly early on in the genesis of this line of research, Salvador et al. (1987) had drawn attention to the difficulty of interpreting results across studies with respect to diverse temporal patterns, stating that ‘...a more detailed study of the temporal course of these variations [T] is required’ (p.13).

One important issue, which subsequently remains unexamined, is the extent to which these variable sampling strategies in general, and the single time-point sampling approach specifically, act to introduce unnecessary error into the results obtained in hormone-competition studies. In order to address questions concerning this issue, as it relates to female competition, it is apposite to draw on tangential evidence and examine what actually occurs to the free fraction of T in females over the course of the day. If it could be demonstrated, for example, that female circulating T not only follows a circadian profile but also exhibits pronounced variability then, given the extremely low levels of T found in female saliva, doubt would be cast on the efficacy of previous attempts to establish meaningful baselines with the use of single time-point sampling; and, additionally, on previous attempts to examine competition changes without having established any baseline. Of single time-point sampling, Riad-Fahmy et al. (1982) note

‘The wide episodic fluctuations in circulating steroid levels make analysis of single samples useful only in screening procedures’ (p.367).

Even though in the fields of endocrinology and clinical chemistry there are a small number early studies on the circadian profiles of female T the results should be viewed with caution, as this is an arena where researchers were limited by the assay technology of the day, invariably measuring total rather than free or bio-available T. In the field of bio-behavioural research only two studies, both of which were conducted by Dabbs (1990; 1991), have examined the circadian variability of female salivary T. These studies are important, occupying a pivotal position in the literature in as much as they are referenced by virtually every single hormone-competition study as justification for a single time-point approach to salivary sampling. Consequently, a close exploration of the methodology and findings of these studies is warranted.

4.2 Circadian Studies

Discussing the findings from their study on circadian and menstrual variation of T, Dabbs and de La Rue (1991) state that ‘Circadian cycles are large enough that they need to be considered in studies of differences among women’ (p.182). Having collected just two samples during the course of a day - one approximately an hour after waking and one just before participants retired to bed for the evening - they reported that morning levels were 80% higher than evening samples. This figure appears surprisingly large, and in evaluating this finding it is worth considering the extent to which methodological shortcomings may limit interpretation of the data from the study. Firstly, with only two samples per day, the collection schedule was extremely limited. In fact, rather than collect samples at predetermined times the participants followed their own schedules. As such, the evening samples had an 85 minute standard deviation. Secondly, the assay performance appeared to be a major limiting factor in placing any confidence in the findings. Not only did the extraction step recover a mere 85% of the original analyte, but within and between-assay coefficients of variation (CV) of 13.4% and 14.8% respectively were excessively high. Moreover, the use of RIA and charcoal extraction in assessing free T provided a lower detection limit of only 6pg/mL. These important and significant limitations would seem to provide a basis for recommending caution when interpreting the results. For the purposes of establishing a protocol for determining baselines in hormone-competition studies, the principal limitation is certainly the lack

of a comprehensive sampling protocol. With only two samples collected there is no possibility of ascertaining the variability in T levels in the intervening period from morning to evening.

Prior to this, Dabbs had already published a paper on the circadian activity of T, in both men and women (Dabbs, 1990). With specific reference to the female participants, he reported data from three independent studies, presenting evidence of circadian activity. One group of participants collected salivary samples at 7am and 10am; a second group collected samples at 7am, 10am, and 10.30am. A final group collected samples at 10am, 4pm, and 10pm. Consistent with the findings in male participants (collected during the same studies) Dabbs reported that female T levels were high in the early morning followed by a drop in the afternoon and early evening. However, these studies also contained serious methodological limitations which, I will argue, render the results uninterpretable. The most egregious error was the inexplicable decision to combine the data from these separate studies which, presumably, involved different participants. In taking this decision Dabbs created a situation whereby a 7am sample for subject A was, potentially, compared against a 10pm sample from subject B in order to determine a circadian profile. I can find no scientific reason to have done this and, given the widespread acknowledgment of often extreme inter-individual variation in levels of T, this process seems indefensible. Moreover, despite earlier in the paper providing several references (c.f. Nicolau et al., 1985; Smals et al., 1976) which pointed to the fluctuation of hormone levels with time of year, the data Dabbs collected, and subsequently combined, were collected at differing times of the year (autumn and spring). Indeed, as Dabbs notes, the study took 2½ years to complete.

In the previous chapter I discussed the challenges faced in accurately measuring female T, particularly the free component as measured in saliva. As was hopefully clear, assay procedures simply cannot be cobbled together or treated as the poor relation in bio-behavioural studies. Again, assay performance and procedures in Dabbs study appear extremely problematic. Firstly, in choosing to analyse aliquots of differing amounts of saliva (from 0.05mL to 0.4mL) Dabbs effectively altered the concentration of T being determined. That he went on to seemingly suggest that the assay was run using under analogue conditions (that is, one standard curve is provided and against which all results are determined). This step alone essentially invalidates the results from

the study. Secondly, the assay performance itself incurred an inter-assay variation in excess of 20% (which is unacceptably high), after recovering only 80-85% of the analyte following extraction. If these events were not problematic enough, Dabbs states 'There were changes in assay materials and in lab technicians and procedures over the course of the studies' (p.84).

And yet, as a direct consequence of the questionable findings from these two research articles which suggest single samples are acceptable in bio-behavioural research, an entire body of literature has adopted a methodological position of single time-point sampling. Having found that 13 out of 74 participants (male participants in study 4) had scores that were more than 3 standard deviation's away from their mean scores Dabbs does, interestingly, and I would argue critically, make the following point, 'This kind of variability should give us pause in working with single measurements from each subject, where it is not possible to recognise a score as deviant from a subject's mean' (1990, p.85). In attempting to determine the validity of utilising single salivary samples for determining baselines there are two issues that arise from this comment. Dabbs appears to be suggesting here that single time-point sampling *can* be problematic, if for no other reason than assay techniques can and do throw up erroneous results that can be extremely difficult to detect unless they are considered in relation to other scores from the same subject. The second issue is still whether circadian variation or episodic fluctuation confounds study design. It is not possible to make a judgement based on this study; firstly because of the serious limitations in the study (design, measurement, methodology, interpretation) but also due to the lack of a comprehensive sampling regimen. It is further worth noting that in his (1991) paper Dabbs completely contradicts his earlier position, stating '...and single measurements are reliable enough for use in behavioral research' (p.815).

The two studies which sought to investigate the T-competition relationship in females (i.e. Bateup et al., 2002 and Mazur, Susman, & Edelbrock, 1997) also adopted this single time-point sampling approach. However, as little is known about detailed daily activity of the free fraction of female T (as measured in saliva), this methodological consideration makes a number of implicit assumptions that may not be justified. Given the paucity of work examining this topic, the different control mechanisms involved, and the different sites of T production between men and women,

it is judicious to investigate circadian profiles of free T in women, in order to clarify the sampling procedures utilised in bio-behavioural studies. Therefore, in order to ascertain the suitability of using single time-point sampling regimens for androgen focused female behavioural research, the present study sought to investigate daily patterns of 'free' T in females in a more comprehensive manner than has previously been undertaken, using assay technology specifically optimised for the measurement of female salivary T.

Having argued that single determination hormone levels may be inappropriate for bio-behavioural studies as a result of the episodic fluctuation, it is reasonable to consider a counter argument; that if T production followed similar pattern on non-consecutive days then single measurements could still be used effectively - assuming the baseline data was collected at the same time on a previous day (i.e. Bateup et al., 2002; Booth et al., 1989). In order to strengthen my argument, that hormone-competition studies involving females should utilise multiple sampling strategies, then it is necessary to demonstrate that on non-consecutive days the reliability between levels of T is low enough to act as an experimental confound. Thus, two studies are reported in this chapter; the first investigating potential circadian activity and episodic fluctuation in T over one day, and the second study examining circadian patterns over the course of two non-consecutive days. This two-day study thus seeks to encompass questions about the methodology of studies which have sought to collect baseline data on the day prior to competition. It further seeks to extend earlier work in this area by utilising a more comprehensive sampling regime. This methodology represents an improvement on previous studies that have either, sampled serum and therefore taken fewer measurements throughout the course of the day (i.e. Dabbs, 1990), or employed small sample sizes.

4.1 Method

4.1.1 Participants

Thirty-six healthy female participants (age range: 18-27yr; mean: 23.5) volunteered to take part in the study, for which each received payment of £5. Payment was made only if study was completed successfully and all samples returned having been stored under the correct conditions. Participants were recruited through advertisements placed with the University's Student Employment Service and around campus. A cursory visual inspection revealed none were hirsute, had serious acne or were overweight. None of the participants were currently administering any form of hormonal medication. Participants indicated what day of their menstrual cycle they were on and all had a history of regular menstrual cycles lasting between 27-34 days.

One significant limitation in a number of previous studies has been the nature of the participants recruited. The human neuro-endocrine system is highly sensitive to a wide range of environmental, biological, and psychosocial stimuli (Sapolsky, 1992) and yet, of the two studies conducted with a female population, one – Mazur, Susman, and Edelbrock (1997) – neglected to enquire if their participants were administering the birth control pill. And yet, as discussed in chapters 2 and 3, by elevating sex hormone binding globulin (SHBG), hormonal medication can critically affect levels of circulating T in an unpredictable manner (Guay, 2002). Consequently, particular attention was paid to exclusion criteria by concentrating on factors that can affect circulating levels of SHBG (see Appendix B for detailed exclusion criteria).

4.1.2 Procedure

Participants met with an experimenter who explained the procedure and obtained written informed consent² (Appendix C). The sampling schedule involved participants collecting salivary samples immediately upon waking, at 9am, and then every 2 hrs throughout the day until 11pm. Participants recorded the time they collected saliva and

² adapted from Sternberg (1993)

were informed that any sample deviating more than 10 minutes from the designated collection time would have to be discarded. A number of authors have identified that, in studying hormonal profiles, compliance with collection protocols can be problematic (Broderick et al., 2004). Consequently, the experimenter emphasised the importance of collecting samples at the allocated time. As participants were highly motivated to participate in the research programme and appeared willing to report salivary collection times outwith their allocated time slots this suggests reasonable compliance with no attempt to deceive (c.f. Dabbs, 1990). In order to take account of environmental and psychosocial effects thought to influence T secretion participants were asked to desist from taking part in any activity that may have modified their T levels. In line with this, they were asked to report any inadvertent event such as bouts of extreme anxiety, alcohol or coffee consumption, elated mood, or sexual activity.

4.1.3 Salivary Collection

For 1hr prior to salivary collection participants were requested not to eat, brush their teeth, drink coffee, or consume dairy products. In order to remove any traces of blood or detritus, on each occasion that participants collected saliva they were instructed to thoroughly rinse their mouths 3 times with water. They waited 3 minutes for the oral environment to normalise and began chewing on a quarter stick of sugar free gum (Wrigley's Orbit Sugar Free: Peppermint). The first mouthful of saliva was discarded in order to remove any unwanted material from the gum which may have potentially interfered with the assay³. All subsequently expectorated saliva was deposited into 20mL polythene universal collection containers upto the 4mL mark. Each sample collection typically took less than 5 minutes, although a small number of participants mentioned they struggled to produce enough saliva. Once collected, samples were stored in a refrigerator overnight and handed back to the investigator the following day for freezing and subsequent assay. Given the low levels of T in female saliva, contamination with haemoglobin has the potential to render analysis meaningless. This is a concern to which we were particularly sensitive. Participants were asked to report if they were aware of any obvious signs of bleeding in their mouths prior to sample collection. Two participants reported evidence of blood in their salivary samples, either from biting cheeks or from dental braces. In these instances, where samples were

³ Larger amounts of gum made assay routinely difficult in the early stages of assay development

obviously and visually contaminated, participants were removed from the analysis entirely, thus accounting for the disparity between participation in the study and the numbers in the analysis.

4.1.4 Hormone Determination

At assay, the frozen salivary samples were removed from the freezer and allowed to thaw. Once thawed, samples were centrifuged at 3500rpm for 10 minutes to break down muco-polysaccharides and aliquoted into a series of smaller eppendorf tubes which were then either re-frozen or centrifuged at 6000rpm for a further 2 minutes prior to assay. Samples were then extracted with ether and T concentrations were determined utilising the 'in-house' ELISA (both these steps are described in detail in chapter 3).

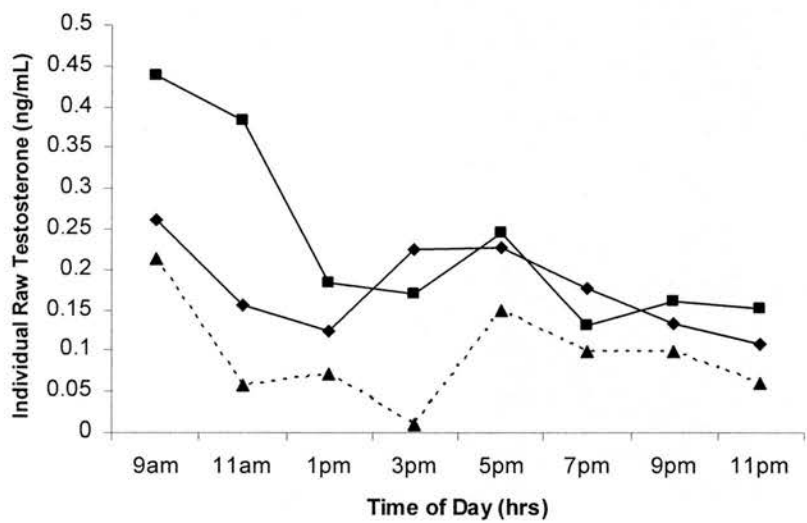
4.1.5 Hormone Data Analysis

Biological data in general, and T data specifically, are often typified by pronounced inter-individual variation. This is likely to be especially true in female studies where menstrual cycle variability is a confounding issue. Using data from three participants participating in the present study Fig. 10 provides an example of the typical spread of raw data.

In order to facilitate graphical representation of the data in a meaningful manner, one way of overcoming the difficulties posed by often extreme inter-individual variation has been to anchor the data to an absolute value. I have subsequently adopted the approach taken by Bateup et al. (2002) and Mazur, Susman, and Edelbrock (1997). That is, the raw data are normalised by taking an individual's highest recorded level of T and dividing the remaining scores by that figure. The data subsequently acquire a figure ranging from above zero to a maximum of 1; the procedure has the added advantage of not distorting the data in any way. Where mean data is reported it is in this normalised format. The large between-subject standard deviations additionally tend to obscure individual trends within group mean data, often leading to non-significant findings. Fig. 10 additionally illustrates that whilst T increases in some individuals it falls in others (between 1pm and 3pm, for example) and these divergent levels ostensibly cancel each other out.

Where authors have collected only a small number of samples (i.e. 1 prior to and 1 post-match) they have tended towards examining any difference in samples as a percentage. Subsequently, in the present study T levels are additionally reported as percentage change from the mean. In the interest of clarity, and where appropriate, raw data scores are also reported. Unless otherwise stated, where sphericity assumptions for the data are not met, within-subjects one-way repeated measures ANOVA tests are reported using the Greenhouse-Geisser correction procedure and ANOVAs are reported on normalised data.

Figure 10 Raw Testosterone Data from Three Participants

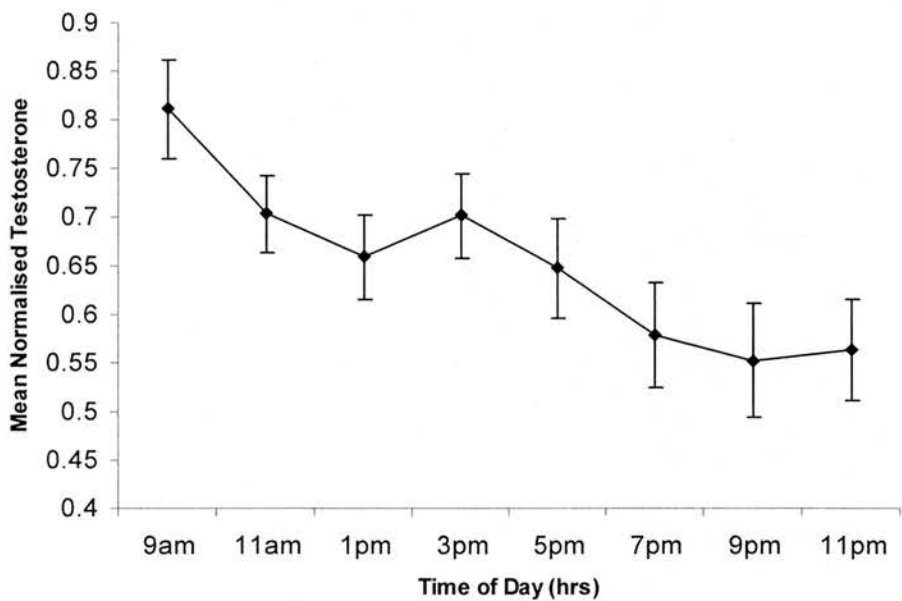


4.2 Results

Consistent with findings from previous studies there was considerable inter-individual variation in T levels. However, Fig. 11 illustrates that within the group mean, female participants demonstrated a clear circadian rhythm in salivary T levels, higher in the morning and lower in the evening. A within-subjects one way repeated measures ANOVA performed on normalised data revealed a significant main effect, e.g. from 9am to 11pm, $F(1, 4.481) = 2.990, p < 0.05$. A subsequent pair-wise Sidak comparison revealed the differences to occur between 9am-9pm and 9am-11pm. It is perhaps of interest to note that whilst ANOVA conducted on normalised data did not reach

significance between 9am and 7pm this actually represents a mean 17% difference. Fig. 12 presents the same data expressed as percentage change from the mean. The degree to which there is inter-individual variability tends to be obscured in this group mean data. Consequently, Table 5 illustrates the magnitude of individual variability around the mean in an illustrative and representative selection of 6 participants. What is particularly revealing is that these fluctuations often occur at a magnitude greater than pre-competition rises in T reported in previous studies. Hence, over and above reporting evidence of circadian activity, an important finding is the highly erratic nature of female T levels throughout the day.

Figure 11 Circadian Female Salivary Testosterone (Mean \pm SEM) $n=34$



As a result of high standard deviations, large inter-individual variability reduces the chances of finding a significant relationship between variables. However, in the current data set, percentage change from mean 9am levels reaches upto 30% (fig. 13) with individual percentage change from mean 9am levels reaching upto 95% (table 6). These individual percentage changes show the greatest variation in the early afternoon (only the 7pm sample displays greater variation) which is the time it has been suggested studies take place in order to account for circadian variation. Compared with 9am levels, T at 9pm were 30.1% lower. The disparity between the findings presented in this study and those guidelines are pursued further in the discussion.

Figure 12 Percentage Difference (\pm SEM) in Salivary Testosterone from Mean Circadian Levels $n=34$

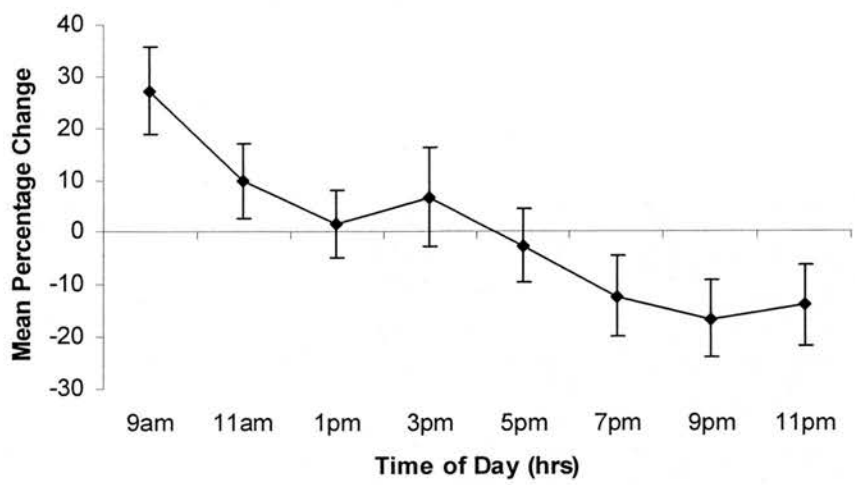


Figure 13 Percentage Difference (Mean \pm SEM) in Raw Salivary Testosterone from 9am Levels $n=34$

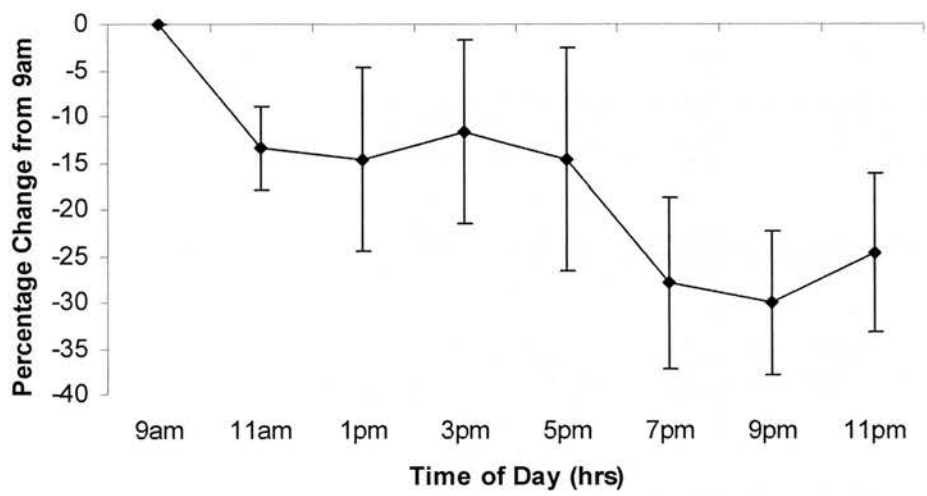


Table 6 Individual % Variability in Testosterone Levels from Circadian Mean

Subject	Time of Day							
	9am	11am	1pm	3pm	5pm	7pm	9pm	11pm
1	25.65	1.10	-16.99	-17.25	-10.27	-0.97	8.07	10.66
2	7.41	-11.11	-24.07	-3.70	-40.74	38.89	11.11	22.22
3	-39.36	-23.40	15.43	47.87	34.57	38.30	-37.23	-36.17
4	60.89	55.67	6.28	-34.69	-22.64	-29.07	-17.82	-18.62
5	-36.91	2.66	11.17	31.60	31.60	29.47	-58.19	-11.38
6	39.80	23.47	19.05	20.75	-56.46	-31.63	-9.18	-5.78

4.3 Introduction

In providing more comprehensive information than has previously been available about the circadian activity of female salivary free T, study 1 illustrated that not only do females experience a distinct circadian rhythm but also, throughout the day, any individual's level of T *can* be highly variable. In a separate, though related, body of research on circadian activity in cortisol (F) Stone et al. (2001) demonstrated that circadian profiles had only modest day-to-day stability; whilst 51% of participants had normal cycles, 17% had depressed (flat) cycles, and 34% had inconsistent cycles across two days. Consequently, given that approximately 25% of the production of T in women emanates from the same site as F it is legitimate to consider if the circadian activity and episodic fluctuation in T occurs similarly across different days. Concerning reliability, Booth, Mazur and Dabbs (1993) reported reliabilities of between $r=0.50$ and $r=0.65$ for T measurements taken at the same time of day in periods ranging from a few days to six years. However, some time earlier Dabbs (1990) had reported reliability levels over the course of two days between $r=0.41$ and $r=0.73$. Because of previously discussed limitations, however, the latter study and findings should be viewed with caution. This disparity does suggest, however, that the reliability between levels of T measured at the same time of day requires further investigation; especially as those hormone-competition studies which have incorporated baseline measurements into their design have invariably collected the baseline sample at least 24hrs prior to the competition. And, in the case of Passelergue and Lac (1999), 3 weeks prior to competition.

4.4 Method

4.4.1 Participants

Twenty-three healthy female participants (age range: 18-23yr; mean: 21.7) with a history of regular menstrual cycles lasting from between 26-34 days volunteered to

participate in the study. All participants also collected samples for the first study and several were participating in a Chief Scientist Office funded project to ascertain circadian and menstrual variation across the female adult life span. None smoked or had drunk alcohol for twenty-four hours prior to the study. None of the participants were administering any form of hormonal medication. Informed consent was obtained from all participants before participation in the study and they were paid £10 for successful completion of the study.

4.4.2 Procedures

The study replicated the instructions for the previous study, the only difference being that participants were asked to collect samples over the course of two non-consecutive days rather than one day.

4.4.3 Salivary Collection and Hormonal Analysis

These procedures were identical to those reported in the previous study.

4.5 Results

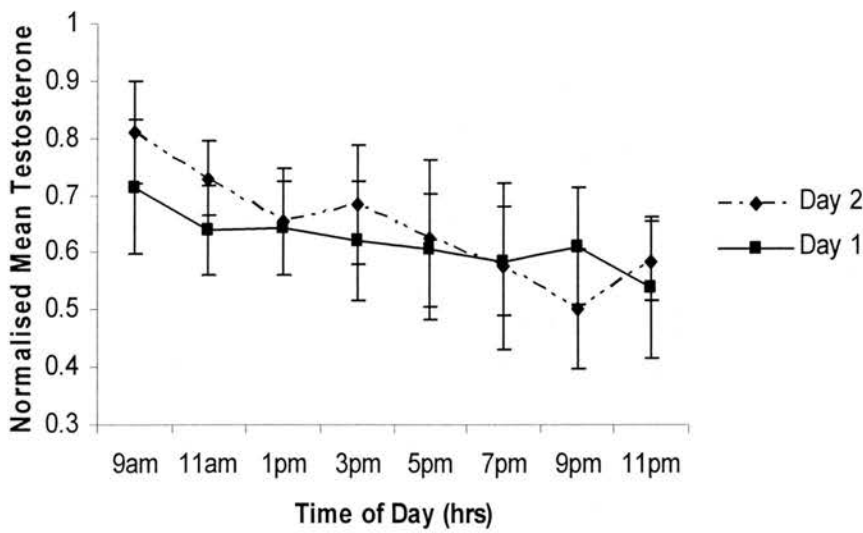
Fig. 14 illustrates that across both days T followed a circadian pattern, with levels higher in the morning and lower in the evenings. The circadian profile appears similar across days. A within subject two-way repeated measures ANOVA with 2 (days) X 8 (time of day) revealed a main effect for time when reporting the Greenhouse-Geisser correction procedure: $F(1, 2.753) = 4.598, p<0.05$. There were no day X time interaction effects: $F(1, 7) = 0.450, p>0.05$. A follow up pairwise comparison using the Sidak procedure revealed the differences to lay between 11pm and 9am, 11am and 1pm. Using a Spearman’s rho correlation, Table 7 shows reliability of mean T at 8 points across the two days. Unless indicated the difference between time points are significant at ≤ 0.05

Table 7 Reliability of Salivary Testosterone on Two Non-Consecutive Days $n=23$

	Time of Day							
	9am	11am	1pm	3pm	5pm	7pm	9pm	11pm
Reliability	0.586	0.567	0.497	0.594	0.359*	0.419*	0.463	0.427*

* n.s.

Figure 14 Normalised Testosterone Over Two Non-Consecutive Days (Mean \pm SEM)
 $n=23$



4.6 Discussion

In utilising single salivary measurements I argue that studies attempting to correlate biological markers with behaviour - and specifically studies attempting to assess the response of T to female competition - have introduced a potentially fatal methodological confound into the design of their studies. This error is occasioned not only by normal circadian activity which, when considered in isolation I propose may be the wrong unit of analysis on which to base the design of hormone-competition studies, but results from the fluctuation evident in un-bound salivary T. The aim of the two studies reported in this chapter was to investigate possible circadian activity of the free component of female T as measured in saliva and to further investigate reliability across two non-consecutive days. By adopting a comprehensive sampling regimen, in which salivary samples were collected every two hours across both one and two days, and by utilising a highly sensitive ‘in-house’ enzyme assay optimised specifically for the measurement of female salivary T, the studies reported here support earlier findings of a circadian profile in female T of approximately 30%. They additionally provide evidence that women experience moment to moment variability of T at levels that call into

question sampling protocols used in the overwhelming majority of T-competition studies. Therefore, these findings address an important gap in the literature on bio-behavioural studies.

In light of the concern expressed over T sampling protocols it is necessary to consider these findings in more detail. Given that for study 1 the only significant differences existed between time points at the beginning and end of the day (9am and 9pm: 9am and 11pm samples) an immediately obvious interpretation of this group mean data would be that single measurements would suffice in behavioural research; collection at one moment in time being likely to yield much the same T level as any other and thus suitable for single time-point sampling in the determination of baseline data. However, the lack of significance owes as much to the high standard deviations produced by extreme inter-individual variation as it does to a lack a difference in levels of T between time points. This only becomes apparent when examining raw data for each participant or the group data as percentage change. For example, whilst the difference between 9am and 3pm failed to reach significance this actually represents a mean 12% drop from 9am levels.

That levels of T in males at waking are high and are subsequently followed by a significant drop in levels is well documented (Bremner et al., 1983). However, there is some suggestion that the extremely high waking levels are a methodological artefact. Swift and James (1982) have suggested that waking samples may represent a concentrated hormone level resulting from evaporation of saliva during the sleeping hours. Consequently once awake and producing saliva normally again the levels return to normal. James and Baxendale (1982) alternatively suggested that the steep fall from waking to early morning was biologically implausible and, in a small scale study, demonstrated that the large variability around early morning samples was a reflection of waking time. If we consider these suggestions and discount the waking samples – as I have done in this study – then the morning levels do not appear less labile than T in the afternoons. Indeed when considering percentage changes from the mean (fig. 13) the greatest degree of lability is actually between 3and 5pm. Thus, further evidence that considering circadian variation in isolation may be the wrong unit of analysis on which to base the design of bio-behavioural studies.

Concerning the development of a suitable T baseline, an important question is: are the levels of episodic fluctuation of a magnitude that might interfere with any attempt to reveal a hormonal response to competition if single time-point sampling is employed? In the most recent study examining hormonal responses to female competition, and which utilised single time-point sampling (one baseline, one pre- and one post-competition sample), Bateup et al. (2002) reported a pre-event rise in T of 24% and a further post-match increases of 49%. This pre-event rise of 24% is only marginally greater than the largest group mean variation reported in this present study.

4.6.1 Sampling Times

In order to control for circadian activity several authors have attempted to collect samples during the course of the afternoon, when T levels are assumed to be less labile than in the mornings. In this regard Booth et al. (1989) state that ‘...it is helpful that matches are played in the afternoon’ (p.558). Echoing this theme, Mazur, Booth, and Dabbs (1992) made explicit reference to their attempts to ‘...ensure that reported effects are not artefacts of...diurnal variation’ (p.72). However, Zitmann and Nieschlag (2001) put forward an alternative account and argue, in their review paper of T and behavioural characteristics, that T samples should be collected during the morning in order to minimise the effects of diurnal variation. Of the two hormone-competition studies involving female participants, Mazur, Susman, and Edelbrock (1997) claimed that by collecting samples between 1pm and 10pm they had effectively controlled for diurnal variation in T. It is my contention that questions exist over the suitability of considering only circadian activity in the design of hormone-competition studies, especially in females where levels of T are often at the absolute limit of assays. I maintain that the episodic fluctuation or random variability across time points reported in the present study occurs at an order of magnitude that ensures comparison between two single points is subject to too much error variance to have any confidence in the findings.

4.6.2 Sources of Error

In considering the validity of single time-point sampling it is as well to note that in bio-behavioural research there are several sources of error, of which random variability and circadian variation are only two. Not least of the remaining sources or error are those occasioned by the performance of the assay itself. Even though the issue of assay performance is not addressed directly in this chapter, it is still worth considering the

effects of cumulative error. If an assay has an intra-assay error of 3-8%, an inter-assay error of 6-15%⁴, combined with the mean error I suggest may come from episodic fluctuation of approximately 15-20% then these sources of error alone would seem to be account for the 24% increase in pre-competition levels of T found by Bateup et al. (2002), where only two samples are being compared. Bear in mind also that this study was naturalistic and was not able to control for the error introduced by physical activity (chapter 2 discussed the effect of not controlling for the physical component of pre-match warm-up, for example). Taking all these very real error sources into account illustrates that an originally impressive increase of 24% can be accounted for and explained away as error variance, with comparative ease. This is not to deny the existence of an effect; considering the errors sources simply draws attention to the fact that there is just no way of ascertaining the validity of an increase of this magnitude.

Similarly, the post match increase of 49% - even if 20-30% of that could be accounted for through sampling and assay error - still leaves an increase in the region of 20%. Where Dabbs (1990) reported T reliability across days of $r=0.41$ and the results from study 2 reported here indicating reliability across days ranging from $r=0.359-0.594$ this further undermines confidence that 49% increase is not substantially an artefact of methodological design. Moreover, given the nature of the study – a female rugby match – there is still the increase resulting from the highly physical component to account for. Indeed in chapter 2, I described this effect in some detail along with the possibility that the samples suffered from blood contamination as a result of the physical nature of the task. It is worthwhile reiterating that whilst I do not suggest that their findings are wrong *per se*, as a result of methodological limitations (i.e. possibly problematic sampling strategies) I believe there is every reason to be cautious in how we interpret the findings from the current female hormone-competition studies.

4.6.3 Mechanisms

If, as I maintain, the random variability of free T is a critically important factor in the design of sampling protocols, then a question of primary importance is: where does this variability emanate from? In males, episodic fluctuation is well documented and the pulsatile release of T derives from the type of feedback loop that controls T –

⁴ These figures represent reasonable and very typical assay performance

gonadotrophin releasing hormone (GNRH) and luteinizing hormone (LH) are released in a fluctuating fashion. However, in females, not only are the feedback loops not well understood but T is produced in three separate locations – adrenal cortex, ovaries, and peripheral conversion of androgen precursors such as androstenedione and DHEA-S. Hence, there exists the possibility that rather than episodic fluctuation – a term which implies a pulsatile mechanism – the variability may result from everyday social interactions and one of the many psychosocial (i.e. stress) or biological factors (i.e. diet) previously demonstrated to impact upon T production. Stress, for example, results in activation of ACTH secretion and therefore testosterone production by the adrenal glands. Whilst an exploration of the molecular control mechanisms falls outside the scope of this thesis, further study is required in order to determine the magnitude of changes brought about by a whole range of everyday situations. Until then, what remains important is that in the vast majority of individuals this variability exists and exists at a magnitude that calls into question the suitability of considering circadian variation only in the design of bio-behavioural studies.

4.7 Summary

In providing a more comprehensive profile of the temporal activity of female salivary free T the two studies reported here deal with an issue not previously considered elsewhere, addressing an important methodological gap in the literature. Study 1 demonstrated that not only is there a small circadian rhythm but random variability can be pronounced. I contend that it is this episodic fluctuation that researchers need to account for and not only circadian activity. Study 2 demonstrated that this temporal profile, whilst similar over two non-consecutive days, produces levels that occasionally have low reliability. Hence, a sample collected at 10am on day 1 may well not correspond particularly closely to a sample collected at 10am on day 2. And yet, this is precisely the type of sampling regimen evident in the hormone-competition literature (e.g. Bateup et al., 2002). Combined, these data strongly support the position that our understanding of hormone-competition relationship in women is seriously hindered by the use of single-time point sampling methodology. A standardised and comprehensive salivary collection protocol, if employed in future studies, would at least allow for a meaningful comparison of chronobiological changes between studies.

Female Hormonal Response to Non-Physical Competition

5. Introduction

The experimental literature on non-human male primates presents a compelling case that testosterone (T) is, in some way – through some as yet poorly understood mechanisms – implicated in achieving a position of social dominance within small group settings, and status processes generally. Moreover, although serving discrete functions, it is suggested that both T and cortisol (F) are implicated in the formation, although not necessarily maintenance, of dominance hierarchies (Bernstein, Gordon, & Rose, 1983). A disparate though substantial body of literature, addressing questions about evolutionary processes, reveals that in a wide range of animal species success in inter-male status encounters confers access to valued resources such as shelter, food, and the increased opportunity of mating with females in oestrus. These are subsequently correlated with health, vigour, and reproductive success (Hall & DeVore, 1965; Richard & Schulman, 1985). Thus, within the context of group living, achieving a position of dominance can have substantial implications for reproductive success in males (Ellis, 1995). Not only that, but among certain primates such as baboons, there are detrimental effects to being subordinate. Research demonstrates that subordinate males may experience significantly higher resting levels of glucocorticoids and they present a hormonal response that is both smaller and slower than in dominant animals (Sapolsky, 1998). Where there is interruption of the HPA axis an individual is exposed to greater risk of infection and illness through chronic stress and impairment of the immune system (Herbert & Cohen, 1993).

This large body of experimental literature, from a whole raft of animal models, including: rodents (Brain, 1990), hens (Allee et al., 1939; Cloutier et al., 1995), fish (Francis, 1983), and non-human primates (Eberhart et al., 1980; Rejeski et al., 1988; Rose et al., 1972), has subsequently attracted substantial attention from researchers

who, in an attempt to confirm a similar endocrine response in humans, have studied the possibility of a dynamic relationship between T, F, and male competition. These interactions have been studied in an assortment of experimental domains including: tennis, wrestling, basketball, judo, chess, and competitive tasks in laboratory settings. Whilst several studies have reported evidence of a dynamic relationship there are those who have been unable to demonstrate that T increases following victory (Elias, 1981; Gladue, Boechler, & McCaul, 1989; Mazur, Susman, & Edelbrock, 1997; Mazur & Lamb, 1980). In respect of these findings, Chapter 2 highlighted that, despite a number of authors neglecting to mention this unpredictability, the nature of these findings are, at best, highly equivocal. In the previous chapter I discussed the possibility that the salivary sampling strategies invariably used in these studies may act as a serious methodological confound; normal steroid metabolism ensuring that T levels fluctuate widely at a magnitude which throws into question the very few findings to date involving females.

As Bateup et al. (2002) note, 'A considerable body of research has focused on hormonal responses to competition; but the participants are nearly always male. Few studies have explored parallel relationships in women' (p.182). Indeed, whilst Bateup et al. (ibid.) recently conducted a naturalistic study of endocrine-behaviour relationships in female rugby players, the only carefully controlled study to date involving female participants focused on males and females competing with same-sex partners in a video game (Mazur et al., 1997). Almost certainly, the reasons for this disproportionate emphasis on the study of males are multifarious; encompassing methodological challenges, gender bias, difficulties in accurately determining T concentrations, and the now increasingly obsolete supposition that T is primarily a male sex-hormone with little role to play in contributing to female behaviour (c.f. Fausto-Sterling, 1992, 2000). The first of these issues – methodology – and specifically the extent to which circadian activity and variability interferes with hormonal sampling protocols was addressed in chapter 4. Secondly, the not inconsequential difficulties associated with accurate determination of free T in saliva were addressed in chapter 3.

And yet, as indicated earlier and discussed in chapter 1, T does exert important physiological and psychological effects, in both men and women. As such it is important to understand these endocrine-behaviour relationships in women, especially

where theoretical and ideological positions concerning gender inequality are maintained upon the findings from biological research, not least of which are hormone-competition studies such as the one reported in this chapter. During the course of the introductory chapter I considered, within the broad framework of evolutionary theory, the extent to which an adequate argument could be constructed to account for women possessing an endocrine response to competition; concluding that, at least theoretically, there were valid reasons to suppose a hormone-competition interaction might exist. Until comparatively recently, for example, it was thought there was little variance in female reproductive success, making dominance contests of little consequence for females. However, there is growing evidence to suggest that competitiveness is useful for females too (Hrdy, 1999; Pusey, Williams, & Goodall, 1997). Moreover, although occasionally it may not appear so, as Kemper (1998) notes, 'We are living in a time of extraordinary social transformation of gender roles, when women seek and attain occupational and political positions that, in men, would be deemed to result *from* or result *in* high T' (p.379). And yet, with an extremely limited evidence base, certain scientists (i.e. Mazur & Booth, 1998) apparently dismiss the possibility of a parallel relationship between T and social dominance in females.

Whilst this thesis primarily focuses on an exploration of the relationship between T and female competition (where it might be seen to serve a function in achieving or maintaining social dominance) and associated issues surrounding the measurement of salivary free T, there are sound conceptual reasons for additionally considering the corticosteroid response to competition. In a range of animal models it too has been associated with status processes, although serving a different function than T (Booth et al., 1989; Sapolsky, 1998). Heightened levels of F have been associated with the experience of stress or anxiety and, in a wide range of animal models, F levels can be seen to rise following defeat and fall following victory in competitive encounters. Indeed, Mazur's biosocial theory of status (1985) primarily conceives of status competitions as an attempt to out-stress an opponent. This specific pattern of endocrine response was thought to demonstrate some of the stress related problems associated with being subordinate. However, additional studies revealed that in a wider range of species it is the experience of victory and not defeat that leads to increased levels of corticosteroid production (Chase, 1980; McGuire et al., 1986). It has been suggested that this contrasting response may occur because the winners of dominance

encounters often face challenges from others soon after gaining high status. Consequently, the response of the corticosteroid system depends very much on a number of factors such as culture, stability of the hierarchy, and generally how detrimental the situation becomes for the subordinate (Sapolsky, 1998). In human males, studies of endocrine response to competition have generally demonstrated the former relationship; that F levels fall in winners and are elevated following defeat.

An additional reason for considering the role of F is that, in males, T is produced primarily in the testes and secondarily in the adrenal cortex. In women, however, production of F and approximately one quarter of total T production occur at the same site – the adrenal cortex (Burger, 2002; Millar & Tyrrell, 1985). This has lead Mazur et al. (1997) to consider the possibility that ‘...some mode of activating the adrenal cortex produces T and C [cortisol] simultaneously. If so, changes in T would parallel changes in C *in women*, because their adrenal cortex is the primary source for both hormones’ (p.318). Indeed, from an evolutionary perspective, this relationship takes on additional significance in females. As Hrdy (1999) notes, stress in a range of species can ‘attenuate menstrual cycles, reduce the period of sexual receptivity, or actually cause reproductive failure’ (p.107).

With these various concerns in mind, a number of questions concerning the relationship between T, F, and female competition are addressed in the present study. These wider questions are considered within the context of the following experimental hypotheses:

1. Females will experience a pre-event rise in testosterone when participating in non-physical dyadic competition over and above that occasioned under baseline conditions
2. Post-competition testosterone will be responsive to outcome (winning vs. losing) in the same way as is suggested in non-human primates and human males
3. Cortisol will co-vary with testosterone in a way that is dissimilar from a standard psychological stress response

During the review of literature I discussed a range of limitations and methodological confounds inherent in previous studies. Limitations such as conceptual difficulties with the win/loss dichotomy, the challenges posed by inaccurate

measurement technique, and the difficulties of interpretation posed by inadequate sampling regimens. In respect of this final point, Salvador et al. (1987) emphasised, ‘...the systematic analysis of the time course of hormone variations in a competitive situation is needed’ (p.13). Despite several studies published since the late eighties, including the two involving female participants, no-one to date has attempted to satisfactorily clarify this important aspect of research design. Additionally, where studies have employed physically demanding experimental conditions (such as tennis or wrestling) they have had to contend with additional methodological challenges. Although the extent and magnitude that hormones are influenced by physically taxing exercise is unclear (there is, for example, no unified response across individuals), it is sufficiently clear that they are a serious confound (Cumming et al., 1987; Gawel et al., 1979; Kuoppasalmi et al., 1976; Sutton & Casey, 1975). In response to concerns over the methodological challenges posed by physical activity Gladue et al. (1989) posed the question, ‘...if subsequent changes in T levels are a by-product of physiological and metabolic excitation of metabolic functioning...can T elevations occur during competition without physical exercise’ (p.411).

Hence, in order to contribute to our understanding of an important topic on which there is currently a paucity of research, the aim of the present study was to ascertain if a dynamic relationship exists between T, F, and non-physical competition in human females. In designing the study I have not only identified a number of methodological limitations in previous studies (chapter 2) but have attempted to address them and, by incorporating improvements into the methodology, limit the extent to which findings from the present study are artefactual products of experimental design.

5.1 Method

5.1.1 Participants

Participants were twenty-two healthy young women aged 18-24yrs (mean 20.4) with self-reported regular menstrual cycles of between 26 and 34 days. None were overweight, smoked, were hirsute, or had administered any form of hormonal medication during the previous nine months. They were recruited through advertisements placed with the University’s Student Employment Service. The study

was approved by the ethics committee of the College of Humanities and Social Sciences, University of Edinburgh.

5.1.2 Study Design

Participants were involved in a non-physical dyadic competition, devised to consist of three elimination rounds and one final, playing the wood block game 'Jenga' (Copyright Hasbro, 2003). Over and above being non-physical, the choice of experimental task was dictated, in part, by the need to minimise possible differences in endocrine response based on skill level. Jones, Hanton, and Swain (1994) note, for example, that skilled or elite performers differ from non-elite performers in as much as they perceive anxiety and stress to be facilitative in their performance; clearly, this difference is likely to fundamentally underpin hormonal changes. For example, in finding that highly ranked male tennis players exhibited consistently lower levels of F than did lower ranked players Booth et al. (1989) suggested that competitors with higher skill levels were likely more adept at managing stress. In response to a question asking 'how would you rate your opponent's skill' all participants answered either 'no idea' or 'about the same'. I take this as evidence that, at least in respect of eliminating skill differences, this was an appropriate experimental task.

As well as minimising disparities in skill level the competition had to be engaging. Mazur and Booth (1998) are at pains to point out, '...in the reciprocal model, as exemplified by the competition studies, T will not rise in response to a challenge when the outcome is a certainty or there is little by way of status or resources at stake' (p.388). This theme is echoed by Bateup et al. (2002) who note that the T-competition relationship is, 'highly contingent on perceptions that gain or loss is at stake' (p.183). Thus, it was necessary to select a non-physical task which, as much as possible, could be played independently of any skill component, whilst still engaging the participants with a robust desire to win. Subsequently, as an incentive to participate in and win this competition, a cash prize of £200 was offered to the overall competition winner. Since the participants invariably had low financial status the large cash prize appeared to be a significant inducement to participate, as adjudged by the sizeable response to recruitment (over 100 young women applied to participate) and acknowledgement of financial reward as the primary reason for participation. In addition to this cash prize, and because of the time requirement for collecting salivary samples and completing

mood and anxiety instruments, each participant was additionally paid £5 for participation, irrespective of outcome for each round. This figure was believed low enough not to interfere with either motivation to win the competition or any sense of disappointment at having lost and being excluded from any further participation.

The study was originally conceived and designed so that experiential effects could be explored. However, approximately halfway through the competition there was significant and unforeseeable participant attrition; one participant fell pregnant, two began taking the pill, two contracted influenza and one left university. It was subsequently impossible for these participants to continue in the competition without compromising the integrity of the data (all apart from the final condition exert an unpredictable effect on SHBG) and so, rather than terminate the study, they were replaced with additional participants. Hence, it was not possible to demonstrate – as had originally been intended – that as participants began to feel increasing pressure with the approaching final that any endocrine response would become more pronounced. Nor indeed to explore questions about the role of prolonged runs of victory on endocrine function. It should be noted that even with these enforced replacements the main aim of the study was not compromised, only this subsidiary question of experiential effects.

5.1.3 Procedure

Participants were initially sent exclusion criteria and a detailed description of the study by email. They subsequently met with the experimenter who confirmed eligibility (exclusion criteria are the same for all studies reported during the thesis; refer to Appendix B), handed them detailed written instructions (Appendix F), and verbally explained the procedure. At this stage written informed consent was obtained (Appendix E).

During each round, 2 competitors played against each other for a maximum of 3 games. The winner of 2 games was subsequently adjudged to be the winner of the match and allowed to proceed to the next round. With 16 participants in the first round, 8 winners progressed through to the next round which was held between 10 and 17 days later. Following the same procedure, 2 winners from that round progressed through to the final. The pair again played 3 games with the winner of 2 games being proclaimed the overall competition winner.

During matches participants were seated opposite each other and separated by a desk. In the middle of the desk was the Jenga tower. The Jenga tower consists of 54 wooden rectangular bricks, in layers of 3, placed at right angles to each other. Each player in turn removes one brick from anywhere below the highest complete storey and places it on the top of the tower, at right angles to the blocks immediately below it. A complete 3-block storey must be completed before starting another. In order to facilitate a sense of pressure a 20 second time limit was placed on the commencement of each move and a loudly ticking timer clock was placed where competitors could both see and hear it. In order to further enhance commitment to succeed in the competition the £200 overall competition cash prize was placed on the table during the competition and the participants were reminded that if they lost they would forfeit their chance to win this money. Following success in 2 matches the winner was declared. The loser was informed they would take no further part in the competition and asked to complete the remaining mood questionnaire and collect the remaining salivary samples.

5.1.4 Procedure for Collection of Salivary Samples

In order to establish comprehensive baseline data for the hormones T and F, participants were asked to collect 6 saliva samples 48hrs prior to competition. When collecting salivary samples participants were required to rinse their mouths thoroughly 3 times with water. The oral environment was then allowed to normalise for 2 mins. Participants then chewed on a quarter stick of sugar-free gum, discarding the first mouthful of saliva which contained cellular debris from the gum. They continued to chew on the gum and deposit saliva into a 20mL universal collection container upto the 4mL mark. Once sufficient saliva had been collected the cap was replaced tightly and samples were returned to the experimenter and placed in a -20°C freezer until ready for assay.

Baseline samples were collected according to the following schedule: 9am and then at 3, 2, 1 hr., 30 mins., and immediately prior to the time of competition. Then, on the day of competition, participants collected salivary samples at times corresponding exactly to the baseline collection schedule. For example, if a competition was to be played at 2pm on a Wednesday then baseline data would be collected at 9am, 11am, midday, 1pm, 1.30pm and, finally, immediately prior to the competition at 2pm on the Monday. Hence, collection of all salivary data was anchored to the start time of each round, which varied among competitors. Following competition, participants collected

an additional 4 samples; starting at 30 mins. after their last game and then every hr. for 3 hrs. Given that nature of the various feedback loops that control steroid metabolism in females, the decision to collect samples from 30mins. post competition reflects the fact that 10mins. – as used in several studies (i.e. Elias, 1981; Mazur et al., 1987; Mazur et al., 1997) – is potentially too short a time in which to see an endocrine response directly related to the competition itself. For example, regarding T metabolism in males, Veldhuis et al. (1987) described an approximately 20 min. delay between LH surge and T secretion. The circadian studies reported in chapter 4 illustrate that over two non-consecutive days reliability of group mean levels of T can be low. Hence, in order to avoid a spurious effect from naturally occurring differences in T levels, samples were also collected at 9am. It was reasoned that if levels varied significantly at 9am then any increase in either T or F between baseline and pre-competition was likely due to natural variation in levels of these hormones rather than reflecting an activational endocrine response in relation to competition.

5.1.5 Hormone Determination

When required for assay, the frozen salivary samples were removed from the freezer and allowed to thaw. Once thawed they were centrifuged at 3500rpm for 10 mins. in order to break down muco-polysaccharides and aliquoted into a series of smaller eppendorf tubes, which were then either re-frozen or assayed immediately. Following ether extraction, salivary T concentration was determined by use of the 'in-house' ELISA described in detail in chapter 3. The assay requires 100 μ l of sample per single determination and all samples were assayed in duplicate. In order to minimise error, all the salivary samples from each subject were assayed in the same plate. Each plate also contained quality control samples which were used to determine intra and inter-assay coefficients of variation which were 2.1% and 6.8% respectively. The mean and standard error T level for all samples in was 0.217 ng/mL \pm 0.009 (range: 0.02-0.41 ng/mL).

Levels of salivary F were determined in Dr. Al-Dujaili's Biological Sciences laboratory at QMUC using the same in-direct assay procedure and using the same ether-extracted salivary samples as those for T determination. Cross-reactivity with cortisone was 1.2%, corticosterone 1.4%, Deoxy-cortisol 1%, testosterone 0.4% and other steroids < 0.5%. Intra and inter assay imprecision values (CV) were 3.65% and 4.74%

respectively. Recovery studies for a range of cortisol levels from 2.6-40.8 ng/mL were 96.8% to 104.7%. The lower detection limit of the assay was 0.03 ng/mL and the standard curve was a highly reproducible $r=0.998$.

5.1.6 Analysis of Hormonal Data

As noted in the previous chapter, T data were typified by pronounced inter-individual variation. Consequently, data from this study were treated with the same normalisation procedure. That is, raw scores were transformed by dividing a participants recorded scores by their highest T level and thus anchoring all values to a maximum of 1. The inter-individual variation has the additional effect of masking individual trends and reduces the likelihood of finding significant differences in group mean data. In order to circumvent this difficulty, and to have some method for examining the magnitude of any changes in levels of T and F in relation to competition phase, percentage change from the mean is additionally reported (c.f. Elias, 1981). This is facilitated by determining the mean of each individual's levels across all time points except the 9am sample and calculating the percentage change for each data point from that mean figure. This has the added benefit of not losing any data or manipulating in too extreme a manner.

The data reported here are for all rounds and participants combined. Hence, participants may be included on more than one occasion (i.e. those winners who progressed through to the following round and who were not excluded from the competition), even though there is no replication in actual data. As a consequence, the data are not strictly un-related. Whilst this approach appears to violate the assumptions of parametric tests it actually need be problematic only in as much as it limits the generalizability of the findings. Bateup et al. (2002) and Booth et al. (1989) utilised a similar approach, treating inter-dependent data as independent. In line with disparate design considerations they treated their statistical analyses somewhat differently. Concerning this procedure, Booth et al. (ibid.) did note 'Our checks of robustness, under various repeated-measures assumptions, convince us that these approximations are not misleading' (p.560).

As with the data from the previous chapter, unless otherwise stated, where sphericity assumptions for the data are not met, within-subjects two-way repeated

measures ANOVA tests are reported using the Greenhouse-Geisser correction procedure and ANOVAs are reported on raw normalised data.

5.1.7 Attitudinal Evaluation and Mood Measurements

Pre-competition: In conjunction with collection of salivary samples, participants were asked to respond to a number of written questions concerning how important they believed the competition to be, how they rated their chances of winning, and what they believed the likelihood of that was. Items were scored with a four point likert scale: 1=not at all, 2=a little, 3=moderately, and 4=extremely. Participants additionally indicated how they rated the skill level of their opponent on a 6 point scale: 1=much worse, 5=a lot better, and 6=no idea. In addition, they were asked to report engaging in any form of visualisation process (reported in Appendix F).

Post-competition: Upon completion of each round, participants were asked how much they believed they had contributed to the outcome, how close they believed the match to be, and how difficult their opponent turned out to be. These items were scored using a 4 point Likert-type scale.

Mood states: Were assessed by means of the original profile of mood states (POMS) which has been demonstrated to be a reliable, valid, and responsive measure of mood status. This self-report instrument comprises 65 questions and provides a score for total mood disturbance as well as for six subscales: Anxiety-Tension, Depression-Dejection, Vigor-Activity, Anger-Hostility, Fatigue-Inertia and Confusion-Bewilderment (McNair, Lorr, & Droppleman, 1971). Subscales have been shown to possess excellent internal consistency and reliability (0.89-0.95) and moderate test-retest reliability (0.65-0.74). For this particular study reliability coefficients determine using Cronbach's Alpha were $\alpha = 0.734$. All subscales except Vigor express negative mood. Where the original version is intended to ascertain how participants felt over the course of the whole day, participants in the present study were asked to respond to how they felt at the time they were completing the instrument.

In order to examine the relationship between competition outcome, levels of T, and mood, optimal scaling regression analyses were conducted on all subscales and total mood disturbance; mood and outcome were treated as independent variables and

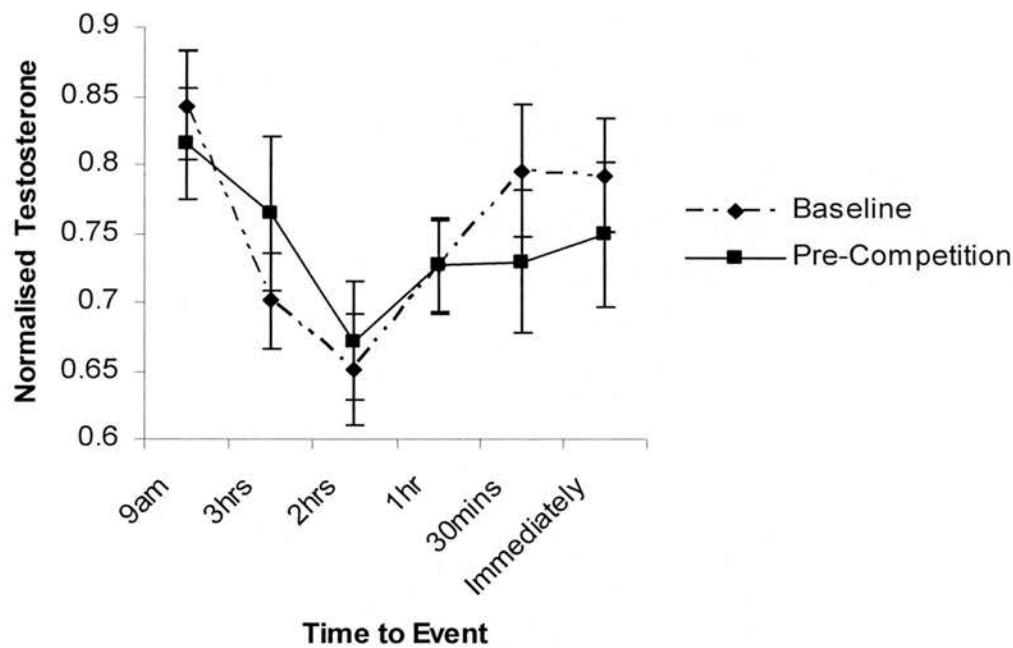
regressed onto T. Results were reported as adjusted R. Standardised beta coefficients were also reported in order to determine the relative contribution of the IV's to the DV. The models tested involved using the change between pre- and post-competition for normalised T and mood.

5.2 Results

5.2.1 Baseline and Pre-Competition Testosterone

Replicating the findings of circadian activity from the two studies in chapter 4, Fig. 15 illustrates that in both the baseline and pre-competition phase, normalised T levels fell from 9am until 2 hrs. prior to competition (this represents a time gap of between 3 and 6hrs.). Under both conditions T then rose until 30 mins. prior to competition. In the final 30 mins., T levels on the day of competition continued to rise, but baseline levels fell marginally.

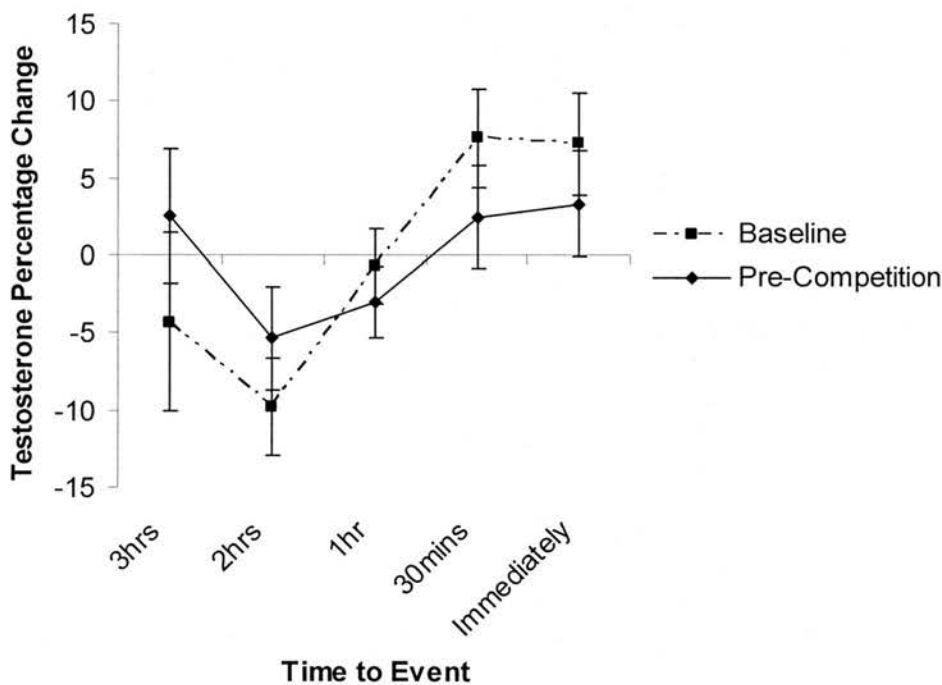
Figure 15 Normalised Testosterone at Baseline and Pre-competition Phase (Mean \pm SEM). $n=30$



A paired samples *t*-test carried out on baseline and pre-competition samples at 9am revealed no significant difference between T levels $t(1, 29) = -.154, p>0.05$. For the baseline and pre-competition phase, a within-subjects repeated measures 2 (competition phase) X 5 (time) two-way ANOVA was performed on the normalised mean data. It revealed no main effect, either for time $F(2.210, 48.614) = 1.454, p>0.05$, or day $F(1, 29) = 0.008, p> 0.05$. There were no interaction effects for either day or time X day.

In order that the reader may better conceptualise the magnitude of these variations Fig. 16 illustrates the same relationship for baseline and pre-competition phase, reported in terms of percentage change from the mean (where the mean excludes the 9am sample).

Figure 16 Percentage Deviation about the Mean on Normalised Testosterone Data at Baseline and Pre-Competition Phase (Mean \pm SEM). *n*=30

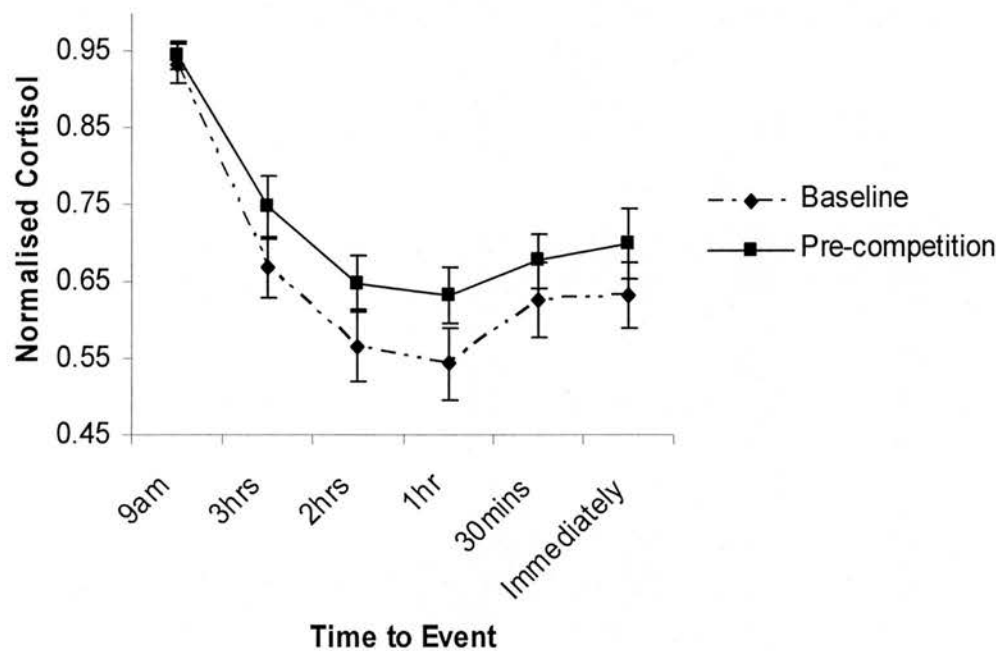


5.2.2 Baseline and Pre-Competition Cortisol

Fig. 17 illustrates that, at baseline and pre-competition, F levels fell from 9am until 1 hour prior to the competition (as with T this actually represents a difference of between 3 and 6 hrs. from 9am). At this point F levels rose until competition. A paired samples *t*-test carried out on baseline and pre-competition samples at 9am revealed no difference

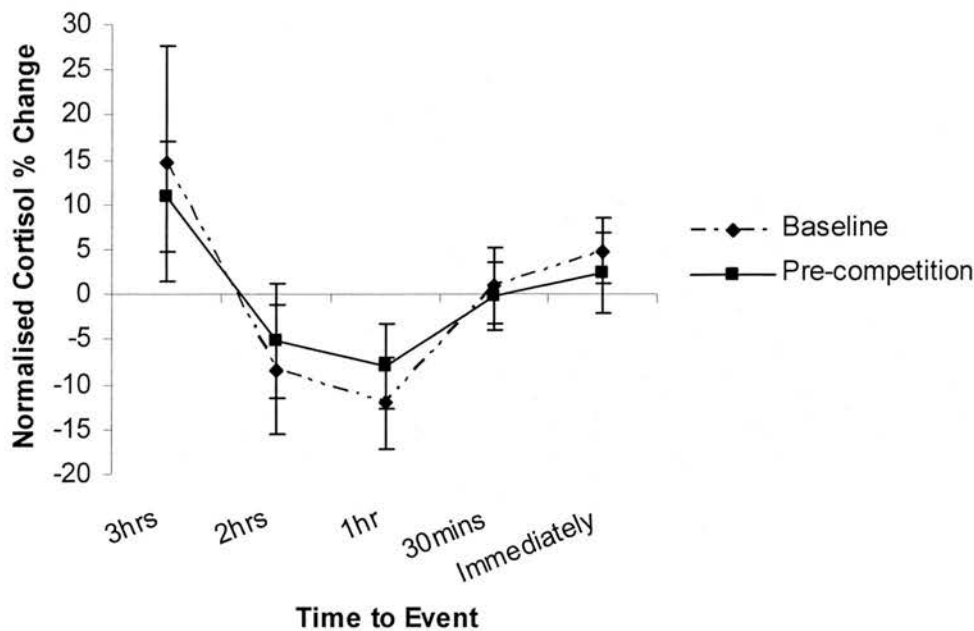
between F levels $t(1, 29) = 0.243, p>0.05$. For illustrative purposes the same data are presented (without inclusion of the 9am sample) as percentage change from the mean in Fig. 18.

Figure 17 Normalised Cortisol (Mean \pm SEM) at Baseline and Pre-Competition Phase.
 $n=30$



At the baseline and pre-competition phases, a within-subjects 2 (competition phase) X 5 (time) two-way repeated measures ANOVA was performed on the normalised mean data, where the mean was established without inclusion of 9am levels. Unlike the T data, the F data passed Mauchley's test for sphericity revealing a main effect for time $F(4, 88) = 3.885, p<0.05$. Follow-up tests using the Sidak calculation revealed the differences were between 3hrs. and 1hr. The main effect for competition phase was non-significant. There were no interaction effects, for either day or day X time. The finding that levels of F rose under the baseline condition, whilst initially surprising, may represent a post-prandial challenge. This finding is considered further in the discussion.

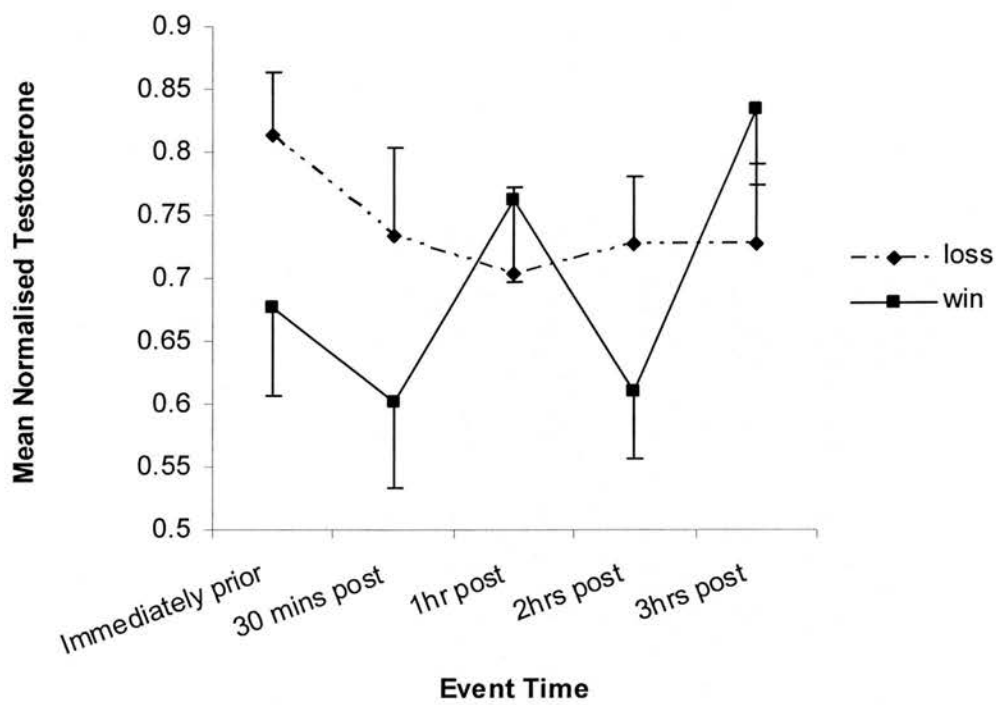
Figure 18 Percentage Change from Mean (Mean \pm SEM) for Normalised Baseline and Pre-Competition Cortisol $n=30$



5.2.3 Post-Competition Testosterone

Mean game length was 16.4 mins. (range: 6-35mins.). T levels in losers fell from immediately pre-competition until 1hr. post-competition. At 2hrs., levels then rose marginally where they remained stable until 3hrs. post-competition. In contrast, T response in winners appeared erratic. T levels fell in the winners for 30mins. before rising at 1hr. post-competition (Fig. 19). At 2hrs. levels had fallen sharply and at 3hrs. levels again rose sharply. For illustrative purposes the same relationship expressed as a percentage from the mean is illustrated in Fig. 21. A within-subjects repeated measures 2 (outcome) X 5 (time) two-way ANOVA performed on the normalised data revealed a main effect for time, having just reached significance $F(4, 48) = 2.547, p \leq 0.05$. Pairwise comparison using the Sidak procedure revealed significant differences between 30 mins and 3 hrs., 1hr. and 2hrs., and finally 2 hrs. and 3 hrs. post-competition. By outcome the simple main effect was non-significant. There were no time X outcome interaction effects.

Figure 19 Normalised Post-Competition Testosterone by Outcome (Mean \pm SEM).
n=14



Given the concerns that single time-point sampling might frustrate interpretation of the data Fig. 20 presents post-competition T at 30mins. and 1hr. separately; illustrating that post-competition results determined by only one salivary sample are susceptible to the vagaries of steroid metabolism. Finally, it is of interest that whilst T levels in losers were higher than winners at 30mins. post-competition they were also higher immediately prior to competition.

Figure 20 Normalised Post-Competition Testosterone by Outcome at 30 mins. and 1 hr. (Mean \pm SEM). $n=14$

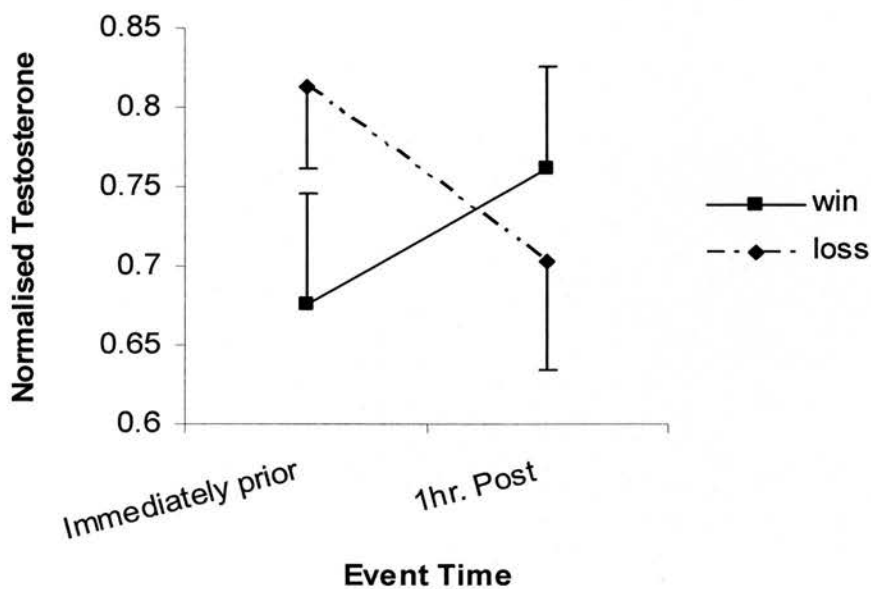
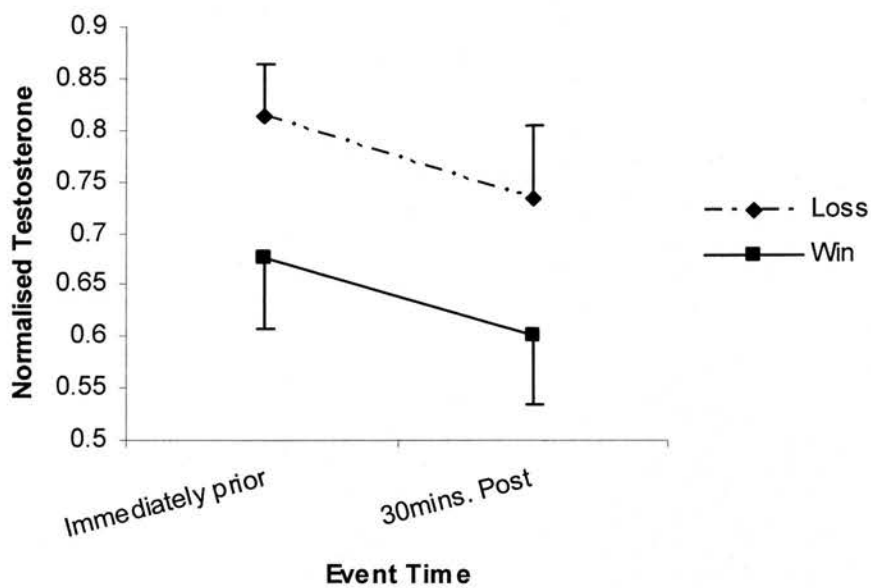
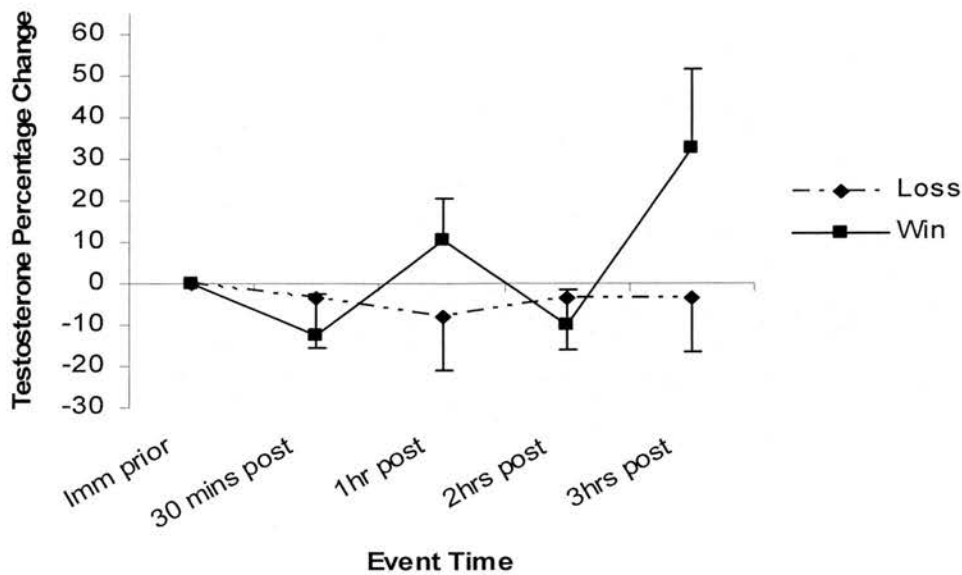


Figure 21 Post-Competition Testosterone Percentage Change from Mean by Outcome
(Mean \pm SEM) $n=14$



5.2.4 Post-Competition Cortisol

Post-competition F levels followed similar patterns by outcome. At 30mins. post-competition, F levels had fallen in both winners and losers, although in winners they fell more sharply. For both outcomes, F levels rose at 1hr., fell at 2hrs., and then rose slightly again at 3hrs. (Fig. 22). For illustrative purposes the same data are presented as percentage change from levels determined immediately prior to competition (Fig. 23). The data passed Mauchley's test for sphericity. A within-subjects repeated measures 2 (outcome) X 5 (time) two-way ANOVA performed on the normalised mean data revealed a main effect for time $F(4, 32) = 2.067, p < 0.05$. Pairwise comparison using Sidak revealed significant differences at immediately prior to competition and both 30 mins. and 2hrs. The main effect by outcome was non-significant. There were no interaction effects for time X outcome $F(4, 32) = 1.230, p > 0.05$. These findings suggest that whilst the winners experienced a reduction in stress following competition the effect was transient. As previous authors have noted, these effects need to be considered in the light of circadian activity; thus, the effects are presumably superimposed upon falling levels.

Figure 22 Post-Competition Cortisol Levels by Outcome (Mean \pm SEM) $n=14$

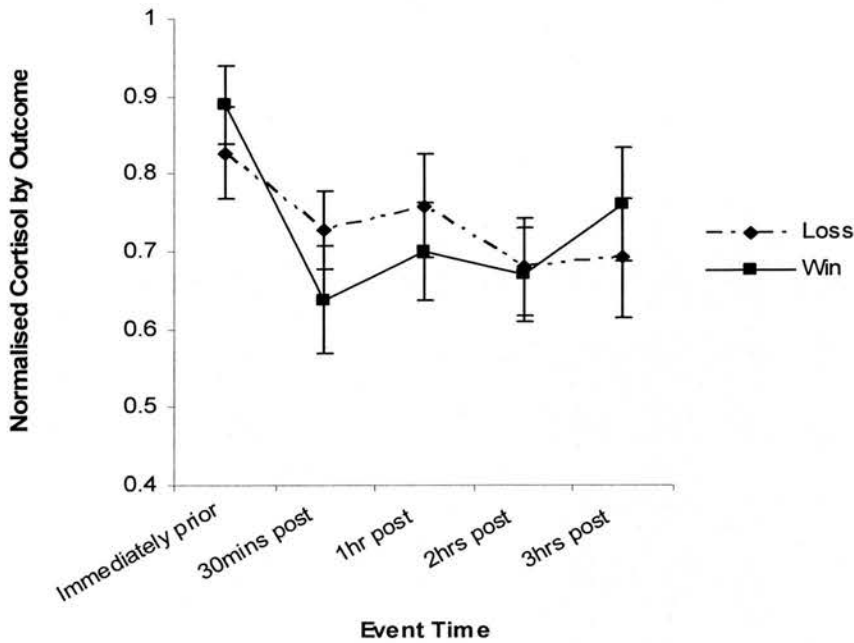
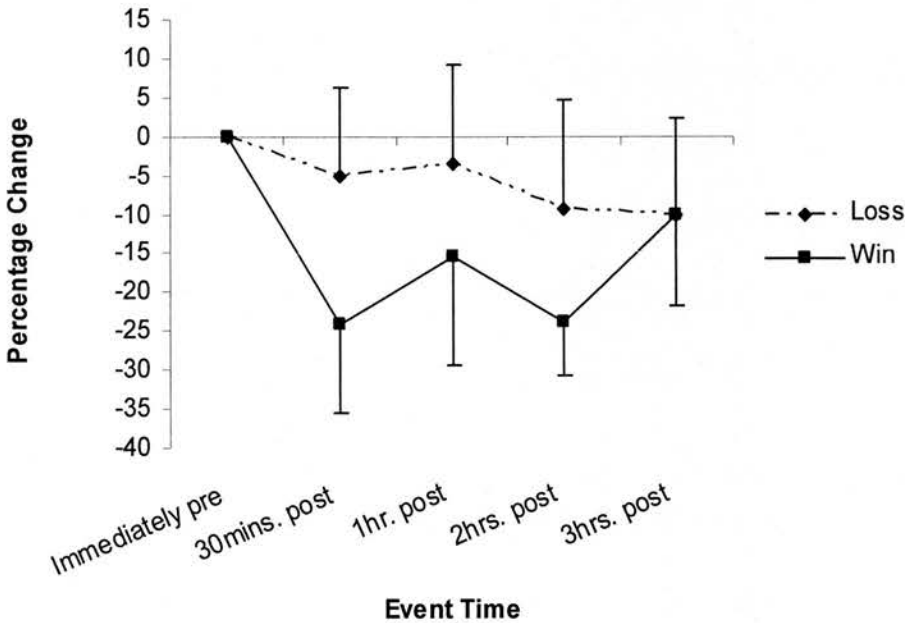


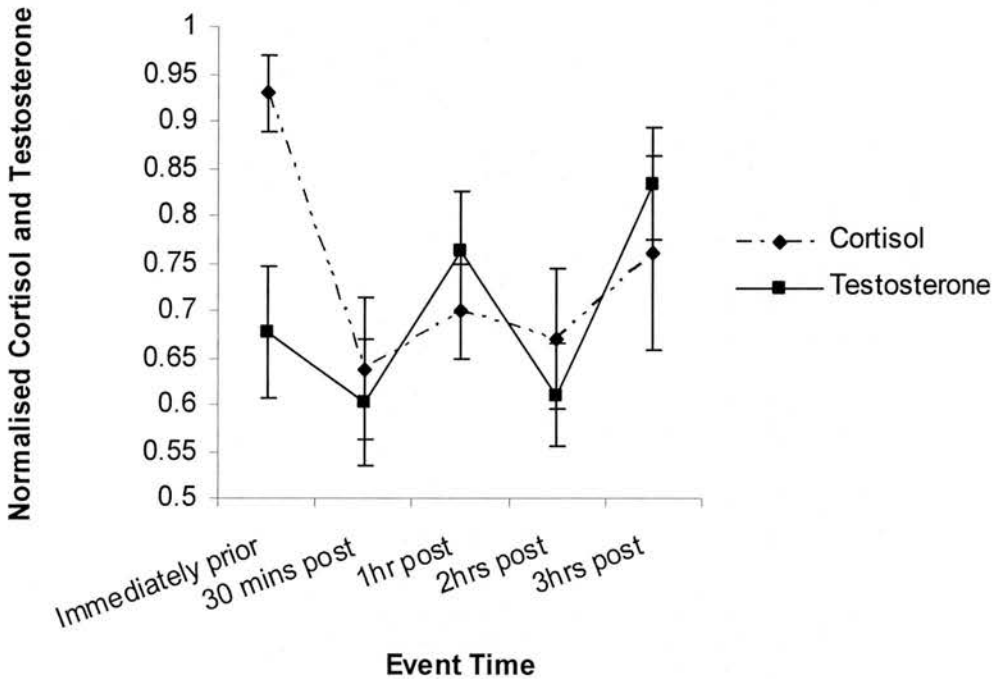
Figure 23 Percentage Change in Post-Competition Cortisol Levels by Outcome (Mean \pm SEM) $n=14$



5.2.5 Cortisol/Testosterone

Fig. 24 illustrates that although differing in relative magnitude, T and F do appear to follow the same profile, changing in unison. A within-subjects repeated measures two-way ANOVA revealed no main effect for hormone $F(1, 13)=0.076, p>0.05$. The simple main effect for time reached significance $F(4, 32) = 3.02, p<0.05$. There were no interaction effects, either by time or by hormonal marker.

Figure 24 Normalised Cortisol and Testosterone (Mean \pm SEM) $n=14$



5.2.6 Testosterone, Outcome and Mood

Table 8 presents the mean and standard errors for POMS data (total mood disturbance and sub-scales) for all participants at baseline, pre-, and post-competition. A one-way ANOVA revealed no simple main effect between baseline, pre-, and post-competition on any of the subscales of the POMS. In considering the relationship between mood, outcome and T it could be argued that the data are not strictly independent, given that the competition involved dyadic encounters. This ambiguity adds to the complexity of the present data set, and it is acknowledged that a number of alternate approaches to analysis of the relationship between mood, outcome and level of T present themselves. However, multiple regression analysis with both mood and outcome as the predictor

variables and T levels as the outcome variable appears to be an appropriate analytical strategy. Analysis involved using the change between pre and post-competition for normalised T (delta T) and mood (delta mood). Because they were categorical, outcome data were dummy coded prior to analysis. The ANOVA output of the regression analyses failed to reach significance. The data were subsequently re-analysed running part and partial correlations and co-linearity diagnostics to ensure that individual data points were not exerting undue leverage on the data set. Under these conditions the ANOVA also failed to reach significance. Therefore, the variance in T is not related to the variance in the IV's. That is, the model does not sufficiently fit the data.

Table 8 Profile of Mood States subscales (Mean \pm SEM) for Baseline, Pre- and Post-Competition ($n = 30$)

Variable	Baseline	Pre-competition	Post-competition
Tension-anxiety	7.34 (0.95)	8.06 (0.99)	6.00 (0.99)
Depression-dejection	7.75 (1.21)	6.59 (1.43)	5.41 (1.21)
Anger-hostility	6.19 (1.15)	5.78 (1.24)	5.63 (1.01)
Vigour-activity	13.47 (1.08)	14.78 (0.97)	13.88 (0.99)
Fatigue-inertia	6.29 (0.77)	5.90 (0.71)	5.56 (0.74)
Confusion-bewilderment	6.75 (0.77)	6.13 (0.83)	4.45 (0.81)
Total mood disturbance	120.8 (4.73)	117.66 (4.74)	114.19 (4.12)

Attitudinal measures concerning how participants felt about the competition and how close they believed the outcome to be ostensible produced a uniform response. Consequently, it was not possible to correlate responses to T levels and so analyses are not reported here. These difficulties are discussed in section 5.4.1.

5.3 Discussion

The main purpose of the present study was to ascertain the existence of any dynamic relationship between T and non-physical competition in women. The comprehensive salivary sampling regimen involved the analysis of approximately 500 samples for both F and T (almost 1200 in total), compared with a typical study involving measurement of between 100 and 200 samples (i.e. 160 female T samples in Mazur et al., 1997). This

allowed for a more thorough examination of the T chronobiological profile than has previously been available.

5.3.1 Baseline and Pre-Competition Testosterone

An example of the complexity of temporal patterning evident in hormone-competition relationships is denoted in the present data, which indicate that although T levels rise prior to competition there is a corresponding rise under the time-matched baseline condition. Whilst T rose in the 2hrs. immediately prior to competition this marginally failed to reach significance. Thus, if the group mean data for pre-competition T are considered in isolation, the null hypothesis is retained. That is, in the pre-competition phase of the study there is no significant increase in levels of T. However, the group mean change over this time period was approximately 9%. A reasonable interpretation might be that, despite not reaching significance, a 9% increase is biologically meaningful; and indeed, this is comparable to the average pre-event rise found in studies with male competitors. The comprehensive nature of the hormonal collection schedule in the present study allowed for a more thorough exploration of chronobiological changes, however. Consequently, when the analysis is modified from a simple examination of the temporal activity of T on the day of competition, and these data are subsequently contrasted with time-matched samples collected at baseline, the null hypothesis is retained; female salivary T is *un*-responsive to competition, over and above any increase occasioned under baseline conditions. Therefore, the data reported here would appear to lend support to the findings of Mazur et al. (1997), who also found women not to experience a rise in T in anticipation of their non-physical task.

Kemper (1990) suggested that where there is evidence of a pre-competition rise in T, this occurs as a result of 'fantasy rehearsal'. Notwithstanding my contention that, due to limitations in sampling protocols and hormonal determination, pre-competition rises in T are far from established in any previous study, there is a substantial body of literature that is concerned with imagery techniques and visualisation procedures used by sports competitors of varying skill levels (c.f. Biddle, 1985; Hale, 1982; Rushall & Lippman, 1997). Moreover, with a small number of animal and human studies illustrating that watching fights (fish, rats) or identifying with a football team (humans) can modify T levels, Kemper's hypothesis that T modification can occur through cognitive means alone is worthy of serious consideration. However, perhaps because of

the nature of the task, no participants reported engaging in a visualisation procedure prior to competition. Speculatively then, perhaps it is the lack of visualisation that accounts, in part, for the lack of a pre-competition rise in T.

Of course, that line of reasoning is predicated on believing a response exists and the experimental conditions simply failed to unearth it. Other explanations for the lack of an anticipatory rise in T might include conceptual arguments closely related to the issue of context. That is, how and where T might be exerting an effect. Bateup et al. (2002) are not alone in having suggested that the mechanisms through which T is implicated in dominance or aggression include making an individual more willing to take risks (Daitzman & Zuckerman, 1980) and improving psychomotor function and co-ordination. In support of the hypothesis that T is implicated in dominance behaviour with an overtly physical component (i.e. in the use of physical violence) Sapolsky (1998) noted that increased T can rapidly elevate rates of cellular metabolism in muscle tissue. Thus, if T is related to activational structures in the body - structures that in some manner prepare the body for physical danger - it would be unsurprising to find that it is not responsive to a primarily psychological competition where the chance of sustaining physical harm is presumably minimal; and the data presented here would appear to support this position. It is worth noting, however, that contrary to Daitzman and Zuckerman (*ibid.*), O'Carroll and Bancroft (1984) found no evidence of T influencing sensation seeking in a placebo-controlled study of eugonadal and hypogonadal men.

In order to develop a sustainable link between androgens and dominance – where non-physical strategies are involved - it would be necessary to ascertain that T also exerted a physiological effect on structures that facilitate competence in these arenas. And indeed, it has been suggested that T leads to improved co-ordination and cognitive function (i.e. increased concentration). Where dominance is competitive and aggressive the influence of T may be more pronounced than where dominance is attained non-physically, i.e. verbally. If this hypothesis is accepted then the challenge becomes one of determining which physiological and psychological structures T impacts upon to facilitate success in any non-physical domain that might lead to success in status achievement.

5.3.2 Baseline and Pre-Competition Cortisol

The initial drop in F levels – from 9am until 2hrs prior to competition – in both the baseline and pre-competition phases was as expected, representing typical (although by no means universal) circadian activity. Whilst levels then rose in the pre-competition condition, in support of the assertion that competitors were taking the competition seriously and were experiencing the event as at least mildly challenging or stressful, levels also rose under the baseline condition. This rise in baseline F was unexpected and, at least initially, challenging to explain. However, a review of the salivary sampling regimen, competition times, and the literature dealing with F measurement, revealed what may well prove to be an experimental artefact. For a variety of seemingly sound methodological reasons the competition was organised to take place in the afternoons. For example, whilst variability in F levels can be marked in the mornings, they are less labile in the afternoons. Moreover, playing matches in the afternoons allowed for a suitable period of time in which to complete the comprehensive pre-competition salivary sampling regimen. However, as the vast majority of the games were played between 1pm and 3pm, it is my feeling that the rise in baseline F may actually represent a well documented, although poorly understood, phenomenon – the post-prandial challenge (Anderson et al., 1987; Gibson et al.1999; Remer, Pietrzik, & Manz, 1998).

5.3.3 Post-Competition Testosterone

Compared against levels determined immediately prior to competition, at 3hrs. post-competition there was a mean 38% increase in T for winners and a mean drop of 4% for losers. However, in winners, the results are erratic and make any simple interpretation of the results challenging. For example, T levels fell for both winners and losers in the samples collected at 30 mins post-competition. At 1hr. T had risen in winners and fallen in losers, and then again at 2hrs. winners levels had fallen again before rising substantially at 3hrs. These findings stand in contrast to a number of physically taxing studies where T levels were elevated following competition, irrespective of outcome (i.e. Bateup et al., 2002).

Similar to the findings of Gonzales-Bono et al., (1999) T by outcome does not appear significantly different (from immediately prior to competition) in the group mean, however. Therefore, similar to the pre-competition phase, these findings appear

to support the second research hypothesis; that female T is un-responsive to competition by outcome. Perusal of the raw data, however, reveals a complex pattern of results, with T in several individuals following the expected relationship (rising T in winners, falling T in losers). Although this is obscured in the group mean data, one interpretation of these findings is that steroid metabolism differs amongst individuals. This would have the effect that individuals differ in the rate at which they display an endocrine response. Hence, whilst some individual winners demonstrated higher levels at 30mins post competition, others demonstrated the same response either at 1, 2, or 3hrs. following the competition, and this accounts for the peculiar variability of levels post competition and the high standard deviation in mean scores. Again, this masking of individual variability by the group mean is a replication of findings in studies such as the one by Booth et al. (1989) who stated 'The effect of win vs loss...masks the failure of the hypothesis in individual cases. Forty-five percent of the winning cases did *not* show a rise [in T]...while 31% of losing cases did' (p.567).

Arguably the most important finding is the patterning of T at 30mins. and 1hr. post-competition (Fig. 20). Here is a clear illustration of the limitations inherent in determination of only 1 post-competition hormone sample. Authors collecting post-competition samples at 30 mins. would have concluded that T levels fall in both winners and losers (i.e. Mazur et al. 1997). If hormone levels were determined at 1hr. however, an entirely different conclusion would have been reached; i.e. T levels rise in winners and fall in losers. Extending this line of reasoning, the present data also indicate that if samples were analysed only at 3hrs. then a different finding again would have been arrived at. If accepted, these findings confirm the wisdom of Salvador et al.'s (2003) observation, that a more thorough examination of temporal patterning is required. Hence, the post-competition findings from the present study, whilst admittedly difficult to interpret from a functional perspective, are suggestive that previous findings based only on one post-competition sample may require re-interpretation.

Whilst not directly a test of Mazur's biosocial theory of status in face to face interaction (1985), the findings reported here do allow for a useful interpretation of it. Mazur tentatively considered that the model would apply equally as well to women although, due to the lack of available research, he acknowledged was unable to draw upon any corroborative evidence. Mazur's theory proposed that following victory,

where T levels are elevated, winners are subsequently more inclined to pursue opportunities to dominate. However, this could presumably only occur if levels were not transient. And yet, even in studies where an increase in levels of T for winners has been reported, these changes do appear transient. The results from the present study reveal that, to the extent female T levels are responsive to outcome, that any changes are highly variable and extremely transient. As such, it is difficult to envisage how this might lead an individual to seek an opportunity to either compete again or to feel compelled to dominate. These findings therefore pose a serious challenge to a significant part of the model; that which assumes increased T following competition are a functional expression a biological urge which leads to striving for dominance.

When all subjects were considered together there appeared to be no difference in any subscales of the POMS at baseline, pre, or post-competition. Where previous studies have been reported as demonstrating that T increases are associated with improvements in mood, these studies have often done something slightly different than correlate T levels with mood. In the present study, making the explicit assumption that mood influences T (rather than T affecting mood), a multiple regression analyses provided the opportunity to examine, in a more sophisticated manner than had previously been attempted, the relative influence of mood and outcome on changes in T levels. The model illustrated that neither mood nor outcome were able to predict levels of T. That outcome was unable to predict T is less surprising than the finding that mood also had little or no predictive power. In chapter 3 I discussed the potential difficulties with using outcome as an objective measure in human hormone-competition studies; further describing the views of a number of authors, who had proposed cognitive factors may be the more salient variable to consider in relation to changing T levels. The findings reported here would appear to partially consolidate those views and provide some indication about where future research programmes might target their efforts; i.e. on psychological factors, over and above mood, relating to competition.

5.3.4 Post-Competition Cortisol

Compared with levels determined immediately prior to the match, F levels in fell in both winners and losers. However, mean levels in winning competitors fell by approximately 25% where in losers this drop was only 4%. Thus, as with the T data there is a tension between tests of statistical significance and the biological consequence

of, for example, a 25% reduction in F levels. However this tension is resolved, the present findings appear supportive of results from several studies of male non-human primates, where losers are reported as being more stressed than are winners, as evidenced by an increased glucocorticoid response. With specific reference to women, the findings consolidate those from Bateup et al. (2002). They stand in contrast, however, to the findings of Suay et al. (1999) who found that F remained uniformly higher in winners than losers. It should be noted that this response, whilst significant at 30 mins. and 2 hrs. post-match, appears transitory. At 3 hrs., levels in both winners and losers had fallen by the same percentage although again, as with T, levels in winners were much more erratic than those of losers.

In support of the final experimental hypothesis, it appears that T did not fall in conjunction with rising levels of F. Indeed, although the relative magnitudes are dissimilar T and F response changed in parallel. Therefore, the findings from the present study appear to support the contention of Mazur et al. (1997), that F and T co-vary in parallel. In many respects this is a surprising finding, given that only 25% of T is produced in the cortex. In an effort to interpret their own finding, Mazur et al. (ibid.) cited a study by Lashansky et al. (1991) in which intravenously injected ACTH leads to increased F but not T. With this in mind, the present findings are highly suggestive that there is something particular about competition that leads to a hormonal response that cannot be simply attributed to situational stress.

5.4 General Discussion

5.4.1 Psychological Variables

Even in non-human primates there is evidence that cognitive appraisal has an important mediating role to play on endocrine effects to competition (Bernstein et al., 1983). As noted earlier, when describing the effect of mood on T levels, several authors have hypothesised that, in humans, cognitive factors may be more important than outcome in the expression of any similar relationship (Salvador et al., 2003). Schultheiss et al. (1999) additionally suggest that personality, rather than simple outcome, may play a part in determining which participants display an endocrine response to competition. Whilst a study of personality variables lays outwith the scope of this thesis, individuals

clearly vary in their motivation and ability to compete for resources, particularly in an ecologically valid setting. Thus, it might be expected that T would be correlated with motivation to win. Unfortunately, the pre-competition questions asking participants how important they felt the competition to be and how they viewed their chances of winning produced a response so uniform that there was simply no variability against which to correlate changes in T. Similarly, where participants were asked to respond to how close they believed the outcome of the match to be, over 98% responded 'very close'. That there was such little variability in response to these attitudinal questions may well be explained as resulting from the contrived nature of the competition; it would appear that everyone was extremely motivated to win the £200, they all believed they had a good chance, and they all believed matches to be extremely close. In more naturalistic studies, such as the one conducted by Bateup et al. (2002), it is perhaps unsurprising that participants differed in their motivation to compete in the event.

Concerning the present study, whilst this uniformity of response was frustrating, it could be argued that the lack of variability represents a high degree of achievement in fulfilling the aims of the experimental design, as differences in skill level and luck did not obviously interfere with an assessment of endocrine response to competition. Speaking with the participants informally about their approach to the competition unearthed some intriguing findings that shed additional light on the hormonal data. Most competitors reported not feeling especially nervous prior to the competition, and this provides at least one explanation for the lack of hormonal response in the pre-competition phase. The second explanation is, of course, that women simply do not possess the endocrinological architecture to respond to the anticipatory demands of competition. However, *immediately* prior to the matches and after having collected their final salivary sample, once seated in front of their opponents with a ticking clock and £200 on the table, many participants reported feeling extremely nervous. It is the nature of determining circulating T levels that this rise in tension immediately pre-competition cannot be revealed by the pre-competition sample. During the matches, tension and anxiety were reported as being exceptionally high, although puzzlingly this is not evident in the post-competition F data. Perhaps additional real-time analysis of catecholamines would capture this emotional response.

In addition, there may be some question concerning the uniformity of response to outcome. That is, similar to Mazur and Lamb (1980), comments made by several participants who had lost following competition revealed that they didn't seem overly concerned. Replies typically involved participants saying they felt 'okay, maybe a bit disappointed, but didn't really expect to win anyway' [finalist]. Given that, pre-competition, participants reported believing they had a good opportunity to win this suggests that participants were cognitively re-structuring their experience following a disappointing outcome. Consequently, this would appear to be an area where further research would benefit our understanding of how psychological variables impact upon endocrine responses to victory or defeat in competitive encounters.

5.4.2 The Importance of Context for the Testosterone-Dominance Relationship

As described in chapter 2, numerous animal studies unequivocally demonstrate that the context in which the T-behaviour relationship is ascertained will impact upon the findings. In some contexts, for example, hormonal influence is less important than environmental factors (Guhl, 1968). As Sapolsky (1997) notes 'Our behavioural biology is usually meaningless outside the context of the social factors and environment in which it occurs' (p.158). This view was echoed by Kemper who stated, 'Women's biological template in one social structural setting may be very different in another setting' (1990, p.13).

Concerning the present study, and indeed all comparable contrived lab based studies (i.e. Gladue et al., 1989), it is acknowledged that the competition is qualitatively different to that experienced under everyday social situations. And here there is some confusion in the literature as to what represents an ecologically valid test of the relationship between T and dominance. In their study of hormonal responses to chess competition, Mazur et al. (1992) note that because their experimental model involved mental rather than physical competition it '...makes the experimental model more similar to normal face-to-face competition' (p.71). In some respects this must be correct. However, other authors have argued that naturalistic competitions (more often than not comprising a physical component) have the advantage of examining endocrine changes in an ecologically valid setting – much closer in environment to the context in which any endocrine response may have evolved. Bateup et al. (2002), in a study involving the game rugby, stated 'This is one of the rare ecologies in which women's

hormonal response to highly physical aggressive competition can be studied' (p.182). Whilst it is undoubtedly important to clarify hormonal responses to physically aggressive competition there may be problems with Bateup et al.'s claims in as much as classification of rugby as aggressive is somewhat simplistic and, I would suggest, unwarranted.

Whilst the competition was very real the participants remained, for the greater part, isolated from each other. Within the context of a larger sporting community, as might be the case for athletes competing in sporting leagues, there is invariably some reputational cost to those competing, and a number of authors (i.e. Cambell, 2004; Hess 2004) clearly describe how reputational attacks have a high degree of salience for females. In a paper considering the inter-relationship between gender, dominance, and stress, Brinkerhoff and Booth (1984) raised intriguing questions about the extent to which individuals realise they are actually taking part in dominance episodes; concluding that 'Much...behaviour...is subtle and its meaning may not always be clear to the participants' (p.161). Thus, whilst the participants in the present study were clearly engaged in a competitive situation it is reasonable to question the extent to which the competition impacted upon their status at all. Whilst an individual might think of gaining status as a result of a victory of having performed well in some event, the fact that participants ostensibly won or lost in an experimental setting that precluded judgement by a wider audience. Hence, whilst complex psychological mechanisms related to interpretation of the situation have been claimed to be more important for hormonal responses than the outcome itself (Salvador et al., 2002) the nature of this task, and in particular the isolation of participants, meant that testing Salvador's conjecture was not really possible. Additional studies need to be constructed in such a manner that there is a greater sense of emotional involvement and subsequently status loss following defeat.

5.5 Limitations

5.5.1 Baselines

In order to interpret these findings correctly it is necessary to consider them within the context of circadian activity. In an attempt to make sense of their unexpected findings,

Booth et al. (1989) considered a phenomenon they termed 'reversal'; the suggestion that irrespective of outcome the natural state for a competitor with high pre-match T would be a return to baseline levels. The studies reported in chapter 4 illustrate that whilst there is evidence of T circadian activity, with a rhythm similar to that found in males, T is seemingly no less labile in the afternoons. Indeed the greatest variations in the two circadian studies occurred between 3pm and 5pm (see chapter 4, Fig. 13). Hence, when attempting to interpret the post-competition hormonal data from the present study, it may indeed be prudent to try and account for circadian profiles as Booth et al. (ibid.) suggested. However, that is perhaps not as straightforward as suggesting that afternoon levels are less reactive or lower than in the mornings; especially not if morning samples are taken after the highly variable waking peak.

Despite employing a more comprehensive baseline sampling schedule than previous studies, the baseline in the current study allowed only for a direct comparison with pre-competition T. As a consequence, the possibility of fluctuating rather than merely falling T levels (resulting from unexplained variability, for example) masking any post-competition hormonal changes remains unaddressed. Given the previously discussed issues regarding sources of error in studies such as these, any evidence for support of a post-competition hormonal response by outcome must, therefore, be viewed cautiously; especially given the erratic nature of winners T levels which render interpretation exasperatingly difficult.

If participants could be persuaded to collect additional salivary samples, with baseline data encompassing the entire period of the competition, this would eradicate the limitation inherent in this present design. In conjunction with my work on circadian activity, the dynamic nature of the T data from the present study provides evidence in support of the suggestion that single time-point sampling is inadequate in hormone-competition studies. It may be, however, that samples collected every hour are unnecessary or, as seems more likely, that a more comprehensive sampling regimen is required in the hour immediately prior to, and post-competition. Whilst it is evident that further work is required to ascertain an optimal sampling regimen for female salivary samples, it is worth noting that the collection of smaller aliquots (i.e. 1mL) could facilitate a salivary sampling protocol every 15 mins. during these critical periods.

5.3.2 Mood Evaluations

The mood measured employed in this study – the original POMS – was, perhaps, not the most appropriate choice of instrument. Whilst extremely well-established, and having the advantage of allowing interpretation of several sub domains of mood, the fact that it was so comprehensive effectively precluded its use across all time points. In their study, Mazur and Lamb (1980) had participants draw continuous graphs of their mood states and, upon further reflection, this may have been a more suitable approach for use with a comprehensive salivary sampling regimen. That said, it should be remembered that whilst T may vary with mood, mood is by no means the only factor affecting levels of T, or indeed *vice versa*. Mazur et al. (1997), for example, found that female winners had improved mood, not increased levels of T. Moreover, questions must be asked regarding the usefulness of any attempt at correlating as precise a measure as salivary T, with a construct as amorphous, changeable, and poorly understood as mood actually is.

5.5.3 Outcome

One of the central issues which increasingly appears to be problematic when interpreting findings from hormone-competition studies is the simplistic nature of objective win/loss dichotomies. Recall from chapter 2 that Elias (1981), in order to justify his choice of experimental paradigm (university wrestlers), claimed that winners and losers were easily identified; a line of reasoning echoed by Booth et al. (1989) who suggested that tennis matches are ‘ideal’ for this type of research, not least because they produce ‘...clear winners and losers’ (p.558). Whilst it may be possible to identify winners and losers in animals more clearly, in humans this is, in fact, more problematic.

People appraise and re-interpret events in their lives in relation to complex psychological schema. The stress literature informs us, for example, that people may employ strategies such as dissonance to re-evaluate their interpretation of a potentially unpleasant event in order to reduce stress (Festinger, 1957). Hence, whilst it *is* possible to objectively identify winners and loser, the interpretation that a competitor places on the outcome of the event requires a more detailed level of investigation – especially given what authors have said about complex psychological processes related to emotional and/or cognitive interpretation of the situation being potentially more important for hormonal responses than the outcome itself (i.e. Salvador et al., 2003;

Schultheiss et al., 1999). Given the uniformity of response to interpretation of the event the present study was not able to capture the complexity of this issue. Indeed, as participants did not know each other or have any emotional investment in the other competitors, the nature of this study precluded a thorough investigation of this theme. Although questions have been asked about the relative importance of the competition for participants no-one, myself included, thought to ask the question at both the beginning and end of the study. Studies taking this area of research forward should seek to incorporate attributional data into their design.

5.6 Summary

Given the wider implications of the subject matter, the possibility that human females possess the neuro-endocrine architecture that might lead to stereotypical endocrine responses to competition, in the same way as is claimed for males, is fascinating. For a variety of reasons there is currently very little work addressing this concept. The two published studies to date - one involving physical competition, the other non-physical – established opposing findings. Given the different domains tested, however, these findings need not necessarily be contradictory. Unfortunately, there are limitations in the respective methodologies that pose serious questions concerning the validity of the data and subsequent findings. Therefore, in addressing an under-researched area and paying particular attention to hormonal measurement, the present study can be seen as making a valuable contribution to our understanding of female hormone-behaviour relationships.

Using the highly sensitive ‘in-house’ enzyme-linked immunosorbant assay optimised for use in this thesis and incorporating a noticeably more comprehensive sampling strategy than employed in previous studies, a quasi-experimental design was employed to study the relationship between T, F, and non-physical female competition. The sampling regimen allowed a more detailed exploration of chronobiological changes in which the complexities of interpreting the hormonal data were revealed. A number of findings emerge from the study. Firstly, whilst group mean T levels rose by around 9% during the pre-competition phase – a figure analogous to findings from studies involving male competitors – this rise also occurred under baseline conditions.

Therefore, the most appropriate interpretation has to be that, under the conditions of the present study, T was not responsive in the pre-competition phase. Post-competition, the differences in levels of T by outcome were extremely revealing. Considered at 30mins. post-competition, T levels for both winners and losers fell. However, at 1hr., levels in losers continued to fall whilst winners experience an increase in levels of approximately 10%; providing support for the assertion of a dynamic relationship between T and competition in human females. Given that these changes did not reach significance, caution is urged when interpreting the data. What this finding does draw attention to, however, is the challenge posed by determining a suitable time frame in which to collect samples. One conclusion could easily be replaced by a contrary one based simply on limited sampling strategies.

Research into androgen-behaviour effects in women is still in its infancy. However, as assay technology continues to develop it is likely to become easier to accurately determine T levels and other important androgenic markers in women. This can only fuel research in this arena; one in which there are still important questions to address. It is my interpretation that we are a long way from being able to assert, with any confidence whatsoever, that hormone levels change in a stereotypical way based on outcome of competition – either in males or females. The methodological limitations of studies already conducted are simply of too great a magnitude. It is my feeling that until additional studies are conducted, using a comprehensive sampling regimen and placing a greater emphasis on psychological variables then, with few exceptions, previous studies (using both male and female participants) provide little meaningful insight into hormone-competition interactions. The following chapter, summarising the thesis, considers various research avenues that would move research in this area forward.

6. Introduction

In considering how testosterone (T) might be implicated in aspects of female social dominance, this thesis has been centrally concerned with examining the relationship between T and non-physical dyadic competition in women. In doing so, questions were raised concerning salivary sampling protocols and measurement techniques. Whilst this thesis makes a contribution to an under researched and poorly understood topic, as it stands amongst very few other studies in its contribution to the field, several more studies similar in design, are required to reinforce confidence in the present findings. With the continuing development of assay procedures that are increasingly capable of accurately determining low levels of free T there is likely to be mounting interest in studying the relationship between biology (especially hormones) and a range of social behaviours in women. It may subsequently be helpful to reflect upon the types of studies that might make a valuable contribute to our understanding of these hormone-behaviour relationships, especially as they relate to our understanding of any biological basis for female social dominance. The following section summarises the content of each chapter and suggests further research agendas.

6.1 Chapter Summaries and Future Research Directions

Chapter 1 discussed the significance of social dominance in human interactions. Adopting, as a starting point, Hrdy's (1999) view that 'The complexity and richness inherent in the social networks female primates forge for themselves has, too often, obscured a vital fact in their lives: that competition among females is central to primate social organisation' (p.96), evidence is introduced to suggest that, despite conflicting and keenly held ideological positions in which contrary arguments are presented,

women may be no less concerned with the formation of hierarchical dominance structures than males. Having established that the view of women as essentially coy, retiring, and unassertive, is effectively the construction of a patriarchal hegemony, the chapter further described the different strategies available for achieving dominance. I introduced evidence to suggest that whilst aggressive competition is one such method, it is by no means the only one. And, at least in humans, aggression is arguably not even the most important, humans being much more concerned with the management of often extremely subtle dominance cues and signals. This can be seen as particularly true of females who, to the extent that they do aggress, generally do so indirectly (Hess, 2004). Acknowledging that competition (physical or non-physical) reflects only one limited facet of social dominance, within a broad framework of evolutionary theory, the feasibility of any hormone-competition relationship in human females was established.

Chapter 2 reviewed the literature concerning hormones and competition. There exists a substantial body of work, from a wide variety of animal models, which attempt to verify the relationship between T and competition (where competition is broadly considered within the contexts of dominance and aggression). The chapter provided a concise description of a selection of findings that illuminate the futility of any direct comparison between male and female studies. An exploration of the literature involving human male subjects summarised the main findings to date and identified a range of methodological limitations which serve to ensure caution should be applied in any interpretation of the findings from these studies. Finally, the two published studies involving female participants were discussed and limitations identified. Consideration of the issues emanating from this chapter were utilised in the design of the circadian studies in chapter 4 and in formulation of the hypotheses generated in chapter 5.

Chapter 3 described the challenges faced in the collection and measurement of T in women, particularly the free fraction as measured in saliva. Discussion of a range of issues revealed that whilst the use of salivary assays is increasing in bio-behavioural research, their use is not always straightforward. Moreover, as commercial assays have been based on protocols derived from measuring male T in serum, their precision and sensitivity ensures they may not be able to accurately determine female free T. Consequently, the optimisation and validation of an extremely sensitive in-house enzyme-linked immunosorbant assay (ELISA) specifically for use in the quantitative

determination of female salivary T was reported. Being extremely sensitive and highly reliable, the performance of the assay was considered satisfactory. Development continues, however, especially with the introduction of newer equipment and in attempting to optimise the ELISA without recourse to a time-consuming extraction step. In order to improve confidence in the results of future hormone-competition studies, especially where female participants are involved and extremely low levels of T are being measured, it would be beneficial if researchers were to assign greater emphasis to subject selection, procedures for collection of salivary samples, and reporting assay performance. Only then can other researchers make informed judgments about the validity of the hormonal data obtained.

Chapter 4 described the daily profile of female salivary free T. In doing so, it was possible to illustrate evidence of circadian activity; with levels approximately 30% higher in the morning than in the late evening. Within the context of this thesis, arguably the more noteworthy findings were that inter-individual disparities in circadian activity are masked within group mean data. And, that intra-individual episodic fluctuation occurs at a magnitude which calls into question the use of single time-point sampling as a viable protocol for determining T baseline data. A second study examining circadian profiles over two non-consecutive days demonstrated similar profiles and reliability at levels that cast additional doubt on the validity of single sampling protocols. Together, these studies address an issue not explicitly considered elsewhere.

Having identified that individual T levels can be highly variable throughout the day, additional studies, using an even more comprehensive sampling strategy would provide a clearer picture of female T biosynthesis. In particular, it would be valuable to know at what rate these fluctuations occur and, given that many environmental and psychological factors impact upon levels of circulating T, the origin of the fluctuations requires elucidation.

Chapter 5 explored the possible existence a dynamic relationship between hormones and female competition, similar to the one widely reported in males. Utilising the findings emanating from chapter 4 as the rationale for devising more comprehensive sampling protocols, a quasi-experimental study was conducted using a non-physical

dyadic experimental task. Whilst there was a clear rise in T prior to competition, it was possible to ascertain that this increase in T was no greater than that occasioned under baseline conditions; a finding that would have been missed had only one sample been collected. The study also revealed that the timing of post-competition sample collection plays a profound role in the results obtained. The comprehensive sampling regimen illustrated that in utilising a single time-point sampling strategy, previous studies have introduced a serious methodological confound into their work which introduces uncertainty about the validity of their findings. Combining a comprehensive sampling regimen and accurate measurement technology this study contributes to our understanding of the relationship between hormones and competition in young healthy women.

In the opening chapter the rationale for thinking about hormonal responses to competition in women was considered. Within the broad framework of evolutionary thinking, the theories of sexual selection (Darwin, 1859) and parental investment (Trivers, 1972) are seen as being central in explaining sex differences in social dominance. Trivers (*ibid.*) points out that the act of reproduction is considerably more expensive in terms of time and energy for women than for men. For example, with pregnancy and lactation women are also relatively limited in the number of offspring they can bear. Hence, the burden of parental investment in humans falls mainly to women. At the most basic level, where there is this greater parental investment it is reasoned this will reduce the likelihood of overt physical aggression; the cost of sustaining serious injury is simply too great.

In line with this, Campbell (1999) argues that women should not be expected to display an endocrine to competition because where they do compete it is for things that have no status value. However, this view stands in stark contrast to researchers like Hrdy (1999) and Cashdan (1995) who strongly argue that females are equally concerned with status and dominance although the mechanisms through which they achieve that are often extremely subtle, both biologically and psychologically. Discussing the social behaviour of monkeys, Rowell (1973) cautioned that the influence of hormones on behaviour is most readily demonstrated in restricted social settings and becomes less and less obvious as the complexity of the social grouping increases.

Where there is overlap in the study of the interaction between androgens, aggression, and dominance, any interpretation of these findings are complicated by the fact that, in both males and females, status hierarchies are not based solely on the threat of physical aggression; other forms of interaction are equally if not more important. Even in non-humans, this is especially evident in the social dominance hierarchies of certain species of baboons and macaques where a female's rank is largely dependent on her ability to obtain and maintain alliances with other females (e.g., Chapais, Girard, & Primi, 1996; Hausfater, Altmann, & Altmann, 1982).

Consequently, any attempt to understand how T is implicated in status behaviour has either, to limit its explanation to dominance where it is achieved aggressively, or extend its explanation to include a wider range of behaviours and cognitive processes. In part, this is where some of the complexity in hormone-social behaviour relationship arises. Given that T operates within a complex web of endocrine parameters and within a larger framework of physiological and psycho-social processes, it seems credible that T exerts its effects permissively. For example, through its effect on serotonin, oxytocin, catecholamines, or indeed on a wide, complex, and inter-connected set of biological processes which lay the biological template for socialization and affiliative behaviour. It has been suggested that oxytocin, for example, facilitates social contact and bonding (Taylor et al., 2000). While both men and women secrete the hormone, oestrogen tends to amplify its effect and T moderates its effect. Moreover, in a study of vervet monkeys, where affiliative behaviours are arguable more important than aggressive behaviour in determining rank (as with humans, generally), Raleigh et al. (1991) found the neurotransmitter serotonin to be highly correlated with dominance rank.

In this light, T may not even be the most suitable biological measure for understanding hormone-competition interactions; markers such as serotonin, DHEA-S, or estradiol may well prove to have more explanatory power. However, given the pervasive nature of T it is entirely appropriate to examine its role in social behaviour, not matter how great or little that may ultimately prove to be. The reality is though, our understanding of how T interacts with genes, other biological markers, and environmental factors to influence social behaviour is still extremely limited. For women this lack of knowledge is even more pronounced. What is most clear is that if T

is implicated in dominance behaviour in women its effect is likely to be extraordinarily small.

6.2 Conclusion

As Simone de Beauvoir (1949) observed, where female behaviour is thought to be more closely regulated by biology than it is in males, this exemplifies an ideological position rather than a factual reality. In my mind it is simply too great a leap to make judgements about the biological foundations of female social dominance behaviour based on hormone-competition studies. It is extremely tempting to simply correlate a hormonal marker, often obtained at one point in time, with as complex a social behaviour as dominance. This urge needs to be resisted. To the best of my knowledge there are no studies which are able to demonstrate temporal link between hormone levels and complex behaviours that could not be equally well explained by psycho-social accounts. Indeed, it is this temporal dislocation which renders any interpretation of hormone-behaviour studies so challenging.

As discussed in chapter 2, any attempt to understand the relationship between hormones and competition requires that attention is paid to the context in which the relationship is studied. As Cashdan notes, 'Our understanding of hormones and dominance in women will require a better understanding of competition among women in domains that are important to them' (1998, p.366). And so, whilst the present study examined the relationship between T and competition under non-physical conditions, there is a clear need for replication. Further studies are warranted that seek to overcome the methodological challenges posed by the effects of exercise on hormones and investigate this relationship under more physically taxing conditions. The distinction between physical and non-physical domains may not necessarily be the most important distinction, however. As discussed in the opening chapter, recent work by Cashdan has cast doubt on previous assumptions that women invariably organise their social groupings around patterns of fraternity. Therefore, research that explores further the manner in which women organise their social groupings, the strategies they employ to do so, and studies about how females act in status-threatening situations are

fundamental to making sense of the kinds of hormone-competition studies that are required in the future.

Perhaps Mazur and Booth (1998) best encapsulate where we find ourselves in understanding the role of T in female social behaviour, 'Research on women is needed to gain even a rudimentary understanding of the relation between hormones and dominance behavior in women' (p.380).

References

- Abraham, G.E. (1974). Ovarian and adrenal contribution to peripheral androgens during the menstrual cycle. *Journal of Clinical Endocrinology and Metabolism*, **39**, 340.
- Ahokoski, O., Virtanen, A., Huupponen, R., Scheinen, H., Salminen, E., Kairisto, V., & Irjala, K. (1998). Biological day-to-day variation and daytime changes in testosterone, follitropin, lutropin and oestradiol-17 β in healthy men. *Clinical Chemistry and Laboratory Medicine*, **36**, 485-491.
- Aldercreutz, H. (1990). *Use of sex hormone and corticosteroid assays in the diagnosis of endocrine disorders in women: Introduction to a round table discussion*. Proceedings of the 4th Symposium on the Analysis of Steroids (pp. 387-395), Pécs, Hungary.
- Al-Dujaili, E.A.S., Denyer, M., Milne, C., Philo, R., Pritchard, J., & Allen, A. (1988) *Serozyme: A manual non-isotopic immunoassay system*. In Immunometric assays; overview & new developments, Proceedings of NEQAS meeting (1-8), Cardiff, Wales.
- Allee, W.C., Collias, N.E., & Lutherman, C.Z., (1939). Modification of the social order in flocks of hens by the injection of testosterone propionate. *Physiological Zoology*, **12**, 412-441.
- Anderson, K.E., Rosner, W., Kahn, M.S. et al. (1987). Diet-hormone interactions: Protein/carbohydrate ratio alters reciprocally the plasma levels of testosterone and cortisol and their respective binding globulins in man. *Life Science*, **40**, 1761-1768.
- Archer, J. (1988). *The behavioural biology of aggression*, Cambridge: Cambridge University Press.
- Archer, J. (1991). The influence of testosterone on human aggression. *British Journal of Psychology*, **82**, 1-28.
- Archer, J. (1994). *Power and Male Violence*. In J. Archer, (Ed.). *Male Violence* (pp. 310-331). London: Routledge.
- Archer, J. (1998). Problems with the concept of dominance and the lack of empirical support for a testosterone-dominance link (commentary). *Behavioral and Brain Sciences*, **21**, 363.
- Atherden, S.M., Corrie, J.E.T., Jones, D.B., Al-Dujaili, E.A.S., & Edwards, R.W. (1985). Development and application of a direct immunoassay for aldosterone in saliva. *Steroids*, **46** (4-5), 845-855.
- Barfield, R.J., Busch, D.E., & Wallen, K. (1972). Gonadal influences on agonistic behavior in the male domestic rat. *Hormones and Behavior*, **3**, 247-259.

- Barrett-Connor (2004). *The role of testosterone in women: Understanding the implications of androgen insufficiency and emerging approaches to care*. 86th Annual Scientific Meeting of the Endocrine Society, Symposium, New Orleans, USA.
- Bateup, H.S., Booth, A., Shirtcliff, E.A., & Granger, D.A. (2002). Testosterone, cortisol, and women's competition. *Evolution and Human Behaviour*, **23**, 181-192.
- Baxendale, P.M., Reed, M.J., & James, V.H.T. (1980). Testosterone in saliva of normal men and its relationship with unbound and total testosterone levels in plasma. *Journal of Endocrinology*, **87**, 46.
- Benton, D. (1992). Hormones and human aggression. In K. Björqvist & P. Niemela (Eds.), *Of mice and women: Aspects of female aggression* (pp. 37-48). Academic Press. Inc.
- Bernhardt, P.C., Dabbs, J.M. Jr., Fielden, J.A., & Lutter, C.D. (1998). Testosterone changes during vicarious experiences of winning and losing among fans at sporting events. *Physiology and Behavior*, **65** (1), 59-62.
- Bernstein, I.S., Rose, R.M., & Gordon, T.P. (1974). Behavioral and environmental events influencing primate testosterone levels. *Journal of Human Evolution*, **3**, 517-525.
- Bernstein, I.S., Gordon, T.P., & Rose, R.M. (1983). *The interaction of hormones, behavior, and social context in nonhuman primates*. In B. B. Svare (Ed.), *Hormones and aggressive behaviour* (pp. 535-562). New York: Plenum Press.
- Besch, N.F., Ruetzel, C.H., Younes, M.A., Huang, N.H., Besch, P.K., & Read G.F. (1982). *Non-chromatographic determination of salivary testosterone in women*: In G.F. Read, D. Riad-Fahmy, R.F Walker, & K. Griffiths (Eds.). Ninth Tenovus Workshop: Immunoassays of Steroids in Saliva. (pp. 221-227), Cardiff, Wales: Alpha Omega.
- Biddle, S.H. (1985). Mental preparation, mental practice and strength tasks: A need for clarification. *Journal of Sports Sciences*, **3**, 67-74.
- Björqvist, K. (2001). Social defeat as a stressor in humans. *Physiology and Behavior*, **73**, 435-442.
- Björqvist, K., Nygren, T., Björklund, A.C. & Björqvist, S.E. (1994). Testosterone intake and aggressiveness: Real effect or anticipation? *Aggressive Behaviour*, **20**, 17-26.
- Bouissou, M-F. (1983). Androgens, aggressive behaviour and social relationships in higher mammals. *Hormone Research*, **18**, 43-61.
- Booth, A., Shelley, G., Mazur, A., Tharp, G., & Kittok, R. (1989). Testosterone, and winning and losing in human competition. *Hormones and Behavior*, **23**, 556-571.
- Booth, A., Mazur, A., & Dabbs Jr., J.M. (1993). Endogenous testosterone and competition: the effect of "fasting". *Steroids*, **58**, 348-350.

Brain, P.F. (1990). *Stress in agonistic contexts in rodents*. In R. Dantzer & R. Zayan (Eds.), *Social Stress in Domestic Animals* (pp. 73-85), Dordrecht: Kluwer.

Braunstein, G.D. (2002). Androgen insufficiency in women: Summary of critical issues. *Fertility and Sterility*, **77** (Suppl. 4), S94-S99.

Bremner, W.J., Vitiello, M.V., & Prinz, P.N. (1983). Loss of circadian rhythmicity in blood testosterone levels with ageing in normal men. *Journal of Clinical Endocrinology and Metabolism*, **56** (6), 1278-1281.

Brewer, M.B., & Caporael, L.R. (1990). Selfish genes vs. selfish people: Sociobiology as origin myth. *Motivation and Emotion*, **14**, 237-243.

Brinkerhoff, D.B. & Booth, A. (1984). Gender, dominance and stress. *Journal of Social Biological Structures*, **7**, 159-177.

Broderick, J.E., Arnold, D., Kudielka, B.M., & Kirschbaum, C. (2004). Salivary cortisol sampling compliance: Comparison of patients and healthy volunteers. *Psychoneuroendocrinology*, **29**, 636-650.

Brown, L.M. (1998). *Raising their voices: The politics of girls anger*. London: Harvard University Press.

Broverman, I.K., Broverman, D.M., Clarkson, F.E., Rosenkrantz, P.S., & Vogel, S.R. (1970). Sex-role stereotypes and clinical judgements of mental health. *Journal of Consulting and Clinical Psychology*, **34**, 1-7.

Bruns, D.E. (2003). The STARD initiative and the reporting of studies of diagnostic accuracy. *Clinical Chemistry*, **49**, 19-22.

Burger, H.G. (2002). Androgen production in women. *Fertility and Sterility*, **77** (4), (Suppl. 4), S3-S5.

Buster, J.E., Meis, P.J., Hobel, C.J., & Marshall, J.R. (1978). Subhourly variability of circulating third trimester maternal steroid concentrations as a source of sampling error. *Journal of Clinical Endocrinology and Metabolism*, **46**, 907.

Camazine, S., Deneubourg, J-L., Franks, N.R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Self-organization in biological systems: Princeton studies in complexity*. Princeton University Press.

Campbell, A. (2004). Female competition: Causes, constraints, content, and contexts. *Journal of Sex Research*, **41**(1), 16-26.

Campbell, A. (1999). Staying alive: evolution, culture and women's intrasexual aggression. *Behavioral and Brain Sciences*, **22**, 203-252.

Campbell, A., Muncer, S., & Odber, J. (1997). Aggression and testosterone: testing a bio-social model. *Aggressive Behavior*, **23**, 229-238.

- Carlström, K. & Rannevik, G. (1990). Sex-hormone-binding globulin and free testosterone assays in clinical diagnosis. Proceedings of the 4th Symposium on the analysis of steroids, Pecs, Hungary.
- Cashdan, E. (1995). Hormones, sex, and status in women. *Hormones and Behavior*, **29**, 354-366.
- Cashdan, E. (1998). Why is testosterone associated with divorce in men? (commentary). *Behavioral and Brain Science*, **21**, 366.
- Cashdan, E. (2004). Why female hierarchies are less stable. *Proceedings of the Annual Meeting of the Human Behavior and Evolution Society*, (Female Competition: Symposium), July, Berlin: Germany.
- Challand, G., Goldie, D., & Landon, J. (1974). Immunoassay in the diagnostic laboratory. *British Medical Bulletin*, **30** (1), 38-43.
- Chambers, K.C. (1998). Target tissue sensitivity, testosterone-social environment interactions, and lattice hierarchies (commentary). *Behavioural and Brain Sciences*, **21**, 366.
- Chapais, B., Girard, M., & Primi, G. (1991). Non-kin alliances and the stability of matrilineal dominance relations in Japanese macaques. *Animal Behavior*, **41**, 481-491.
- Chard, T. & Martin, M.J. (1974). *Posterior pituitary peptides*. Radioimmunoassay and saturation analysis: Basic principles and theory. *British Medical Bulletin*, **30** (1), 74-79.
- Chase, I.D. (1974). Models of hierarchy formation in animal societies. *Behavioral Science*, **19**, 374-382.
- Chase, I.D. (1980). Social process and hierarchy formation in small groups: A comparative perspective. *American Sociological Review*, **45**, 905-925.
- Chatterton, R.T. Jr., Vogelsong, K.M., Lu, Yu-Cai., & Hudgens, G.A. (1997). Hormonal responses to psychological stress in men preparing for skydiving. *Journal of Clinical Endocrinology and Metabolism*, **82** (8), 2503-2509.
- Cheng, R.W., Reed, M.J., & James, V.H.T. (1986). Plasma free testosterone: equilibrium dialysis vs. direct radioimmunoassay. *Clinical Chemistry*, **32**, 1411.
- Chiras, D.D. (1999). *Human biology: Health, homeostasis, and the environment*. Sudbury, MA: Jones & Bartlett publishers.
- Christenson, S.E. & Breedlove, S.M. (1998). Seductive allure of dichotomies (commentary). *Behavioural and Brain Sciences*, **21**, 367.
- Clements, A.D. & Parker, R.C. (1998). The relationship between salivary cortisol concentrations in frozen versus mailed samples. *Psychoneuroendocrinology*, **23** (6), 613-616.

- Cloutier, S., Beaugrand, J.P., & Laguë, P.C. (1995). The effect of prior victory or defeat in the same site as that of subsequent encounter on the determination of dyadic dominance in the domestic hen. *Behavioural Processes* **34**, 293-298.
- Clow, A., Thorn, L., Evans, P., & Hucklebridge, F. (2004). The awakening cortisol response: Methodological issues and significance. *Stress*, **7** (1), 29-37.
- Cohen, D., Nisbett, R.E., Bowdle, B.F., & Schwartz, N. (1996). Insult, aggression, and the southern culture of honor: An 'experimental ethnography', *Journal of Personality and Social Psychology*, **70** (5), 945-960.
- Cosmides, L., & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge, MA: MIT Press.
- Cumming, D.C., Wall, S.R., Galbraith, M.A., & Belcastro, A.N. (1987). Reproductive hormone responses to resistance exercise. *Medicine and Science in Sports and Exercise*, **19**, 234-238.
- Dabbs, J.M. Jr. (1990). Salivary testosterone measurements: reliability across hours, days, and weeks. *Physiology and Behavior*, **48**, 83-86.
- Dabbs, J.M. Jr. (1991). Saliva testosterone measurements: collecting, storing, and mailing saliva samples. *Physiology and Behavior* **49** (4), 815-17.
- Dabbs, J.M. Jr. (1993). *Salivary testosterone measurements in behavioral studies*. In D. Malamud, & L.A. Tabak. (Eds.). *Saliva as a diagnostic fluid. Annals of the New York Academy of Sciences*, Vol. 694. (pp. 177-183). New York, NY: New York Academy of Sciences.
- Dabbs, J.M. Jr. (1998). Testosterone and the concept of dominance. *Behavioral and Brain Sciences*, **21**, 370.
- Dabbs, J.M. Jr. (2000). *Heroes rogues and lovers: Testosterone and behavior*. New York: McGraw-Hill.
- Dabbs, J.M. Jr. & de la Rue, D. (1991). Salivary Testosterone measurements among women: relative magnitude of circadian and menstrual cycles. *Hormone Research*, **35** (5), 182-184.
- Dabbs, J.M. Jr., Strong, R., & Milun, R. (1997). Exploring the mind of testosterone: a beeper study. *Journal of Research in Personality*, **31**, 557-587.
- Darwin, C.R. (1859). *The origin of species*. London: John Murray.
- Daitzman, R. & Zuckerman. M. (1980). Disinhibitory sensation seeking, personality and gonadal hormones. *Personality and Individual Differences*, **1**, 103-110.
- Davis, S. & Tran, J. (2001). What are "normal" testosterone levels for women?. *Journal of Clinical Endocrinology and Metabolism*, **86** (4), 1842.

- Davis, S. & Tran, J. (2001). Testosterone influences libido and well being in women. *Trends in Endocrinology & Metabolism*, **12** (1), 33-38.
- de Beauvoir, S. (1949). *The Second Sex (La Deuxieme Sexe)*. Trans. H.M. Parshley (1997). London: Vintage.
- de Waal, F.B.M. (1986). The integration of dominance and social bonding in primates. *The Quarterly Review of Biology*, **61** (4), 459-479.
- de Waal, F. (2001). *The Ape and the Sushi Master: Cultural reflections of a primatologist*. New York: Basic Books.
- Decker, S.A. (2000). Salivary cortisol and social status among dominican men. *Hormones and Behavior*, **38**, 29-38.
- Demers, L.M. (1995). *Biochemistry and laboratory measurement of androgens in women*. In G.P. Redmond (Ed.). Androgenic disorders. New York: Raven Press.
- Dessypris, A., Kuoppasalmi, K., & Alderkreutz, H. (1976). Plasma, cortisol, testosterone, androstenedione and luteinizing hormone (LH) in a non-competitive marathon run. *Journal of Steroid Biochemistry*, **7**, 33-37.
- Dixson, A.F. (1980). Androgens and aggressive behavior in primates: A review. *Aggressive Behavior*, **6**, 37-67.
- Dunbar, R. I. M. (1988). *Primate social systems*. Ithaca, NY: Cornell University Press.
- Eagly, A. & Steffen, V.J. (1986). Gender and aggressive behaviour: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, **100**, 309-330.
- Eberhart, J.A., Keverne, E.B., & Meller, R.E. (1980). Social influences on plasma testosterone levels in male talapoin monkeys. *Hormones and Behavior*, **14**, 247-266.
- Edwards, R. (1985). *Immunoassay: An introduction*. London: William Heinmann.
- Edwards, D.A. & Herndon, J. (1970). Neonatal estrogen stimulation and aggressive behaviour in female mice. *Physiology and Behavior*, **5**, 993-999.
- Ehrenkranz, J., Bliss, E., & Sheard, M.H. (1974). Plasma testosterone: Correlation with aggressive behaviour and social dominance in man. *Psychosomatic Medicine*, **36**, 469-475.
- Eisenberg, L. (1972). The human nature of human nature. *Science*, New Series, **176**, 4031, 123-128.
- Eisenberg, J. F., Muckenhirn, N. A., & Rudran, R. (1972), *Science*, New Series, **176**, 4037, 863-874.
- Ekins, R.P. (1974). Radioimmunoassay and saturation analysis: Basic principles and theory. *British Medical Bulletin*, **30** (1), 1-11.

- Elias, M. (1981). Serum cortisol, testosterone, and testosterone-binding globulin responses to competitive fighting in human males. *Aggressive Behavior*, **7**, 215-224.
- Ellis, L. (1995). Dominance and reproductive success among nonhuman animals. *Ethology and Sociobiology*, **16**, 257-333.
- Ellyson, S.L. & Dovidio, J.F. (1985). *Power, dominance, and nonverbal behaviour: Basic concepts and issues*. In Ellyson, S.L. & Dovidio, J.F. (Eds.) *Power, dominance, and nonverbal behaviour* (pp. 1-27). New York: Springer.
- Eubank, M.R., Collins, D., Lovell, G., Dorling, D., & Talbot, S. (1997) Individual temporal differences in pre-competition anxiety and hormonal concentration. *Personality and Individual Differences*, **23** (6), 1031-1039.
- Fausto-Sterling, A. (1992). *Myths of Gender: Biological theories about women and men*. New York: Basic Books.
- Fausto-Sterling, A. (2000). *Sexing the Body: Gender politics and the construction of sexuality*. New York: Basic Books.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row Peterson.
- Filaire, E., Maso, F., Sagnol, M., Ferrand, C., & Lac, G. (2001). Anxiety, hormonal responses, and coping during a judo competition. *Aggressive Behavior*, **27**, 55-63.
- Fiske, D. W. (1971). *Measuring the concepts of human personality*, Chicago: Aldine.
- Flinn, M., Baerwald, C., Decker, S., & England, B. (1998). Evolutionary functions of neuroendocrine response to social environment (commentary). *Behavioral and Brain Sciences*, **21**, 372.
- Francis, R.C. (1983) Experimental effects on agonistic behaviour in the paradise fish, *Macropodus opercularis*. *Behavior*, **85**, 292-313
- Frey, H. (1982). The endocrine response to physical exercise. *Scandinavian Journal of Social Medicine*, Suppl. **1**, 29, 71-75.
- Frodi, A., Macaulay, J., & Thome, P.R. (1977). Are women always less aggressive than men? A review of the experimental literature. *Psychological Bulletin*, **84**, 634-660.
- Gawel, M.J., Park, D.M., Alaghband-Zadeh, J., & Rose, F.C. (1979). Exercise and hormonal secretion. *Postgraduate Medical Journal*, **55**, 373-376.
- Gibson, EL., Checkley, S., Papadopoulos, A., Poon, L., et al. (1999) Increased Salivary Cortisol Reliably Induced by a Protein-Rich Midday Meal. *Psychosomatic Medicine*, **61**, 214-224
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

- Gladue, B.A., Boechler, M., & McCaul, K.D. (1989). Hormonal response to competition in human males. *Aggressive Behavior*, **15**, 409-422.
- Goldstein, J.S. (2001). *War and gender: How gender shapes the war system and vice versa*. Cambridge, MA: Cambridge University Press.
- Goldzieher, J.W., Dozier, T.S., Smith, K.D., & Steinberger, E. (1976). Improving the diagnostic reliability of rapidly fluctuating plasma hormone levels by optimised multiple-sampling techniques. *Journal of Clinical Endocrinology and Metabolism*, **43**, 824.
- Gonzalez-Bono, E., Salvador, A., Serrano, M.A., & Ricarte, J. (1999). Testosterone, cortisol, and mood in a sports team competition. *Hormones and Behavior*, **35**, 55-62.
- Granger, D.A., Schwartz, E.B., Booth, A., & Arentz, M. (1999). Salivary testosterone determination in studies of child health and development. *Hormones and Behavior*, **35**, 18-27.
- Granger, D.A., Shirtcliff E.A., Booth, A., Kivlighan, K.T., & Schwartz, E.B. (2004). The Trouble with Salivary Testosterone. *Psychoneuroendocrinology* **29**, 1229-1240.
- Grant, V.J. (1994). Maternal dominance and the conception of sons. *Journal of Medical Psychology*, **67**, 343-351.
- Grant, V.J. & France, J.T. (2001). Dominance and testosterone in women. *Biological Psychology*, **58**, 41-47.
- Griffin, J.E. & Wilson, J.D. (1985). *Disorders of the testis and male reproductive tract*. In J.D. Wilson & D.W. Foster (Eds.) *Williams Textbook of Endocrinology*, (7th Edition pp. 259-311), Philadelphia: WB Saunders.
- Grosz, E. (1994). *Volatile Bodies: Toward a Corporeal Feminism*. Bloomington: Indiana University Press.
- Guay, A.T. (2001). Advances in the management of androgen deficiency in women. *Medical Aspects of Human Sexuality*, **1**, 32-8.
- Guay, A.T. (2002). Screening for androgen deficiency in women: Methodological and interpretative issues: The Princeton consensus statement on definition, classification, and assessment. *Fertility and Sterility*, **4** (Suppl. 4) S83-88.
- Guhl A.M. (1968). Social inertia and social stability in chickens. *Animal Behavior*, **16** (2), 219-32.
- Hale, B.D. (1982). The effects of internal and external imagery on muscular and ocular concomitants. *Journal of Sport Psychology*, **4**, 379-387.
- Hall, P. (1988). In Knobil, E. & J.D. Neill (Eds.). *Physiology of Reproduction*. Raven Press.

- Hall, K.R.L. & deVore, I. (1965). *Baboon Social Behavior*. In I. de Vore. (Ed.) *Primate social behavior: Field studies of monkeys and apes* (pp. 53-110). New York: Holt, Rinehart and Winston.
- Hammond, G.L., Nisker, J.A., Jones, L.A., & Siiteri, P.K. (1980). Estimation of the percentage of free steroid in undiluted serum by centrifugal ultrafiltration-dialysis. *Journal of Biological Chemistry*, **255** (11), 5023-5026.
- Harris, J.A. (1999). Review and methodological considerations in research on testosterone and aggression. *Aggression and Violent Behavior*, **4** (3), 273-291.
- Hausfater, G., Altmann, J., & Altmann, S.A. (1982). Long-term consistency of dominance relations among female baboons (*Papio cynocephalus*). *Science* **217**, 752-755.
- Hawley, P.H. (1999). The ontogenesis of social dominance: a strategy-based evolutionary perspective. *Developmental Review*, **19**, 97-132.
- Herbert, T.B. & Cohen, S. (1993). Stress and immunity in humans: A meta-analytic review. *Psychosomatic Medicine*, **55**, 364-379.
- Heritage, A., Strumpf, W., Sar, M., & Grant, L. (1980). Brainstem catecholamine neurons are target sites for sex steroid hormones. *Science*, **207** (21), 1377-1379.
- Herold, D.A. & Fitzgerald, R.L. (2003). Immunoassays for testosterone in women: better than a guess? *Clinical Chemistry*, **49** (8), 1250-1251.
- Hess, N.C. (2004). Evolved sex differences in aggression types: Inclinations towards physical vs. reputational attacks. *Proceedings of the Annual Meeting of the Human Behavior and Evolution Society*, (Female Competition: Symposium), July, Berlin: Germany.
- Hines, M. (1998). Adult testosterone levels have little or no influence on dominance in men (commentary). *Behavioral and Brain Sciences*, **21**, 377.
- Hofman, L.F. (2001). Human saliva as a diagnostic specimen. *Journal of Nutrition*, **131** (5), 1621S-1625S.
- Honk, J. V., Tuiten, A., Verbaten, R., Van den Hout, M., Koppeschaar, H., Thijssen, J., & de Haan, E. (1999). Correlation's among salivary testosterone, mood, and selective attention to threat in humans. *Hormones and Behavior*, **36**, 17-24.
- Hrdy, S. (1999). *The Woman That Never Evolved*. (revised edition). Cambridge, MA: Harvard University Press.
- Hubl, W., Taubert, H., Freymann, E., Meissner, D., Stahl, F., & Dorner, G. (1984). A sensitive direct enzyme immunoassay for cortisol in plasma and saliva. *Experimental Clinical Endocrinology*, **84**, 63-70.

- James, V.H.T. & Baxendale P.M. (1982) *Androgens in Saliva*. In G.F. Read, D. Riad-Fahmy, R.F. Walker, & K. Griffiths (Eds.). Ninth Tenovus Workshop: Immunoassays of Steroids in Saliva. (pp. 193-201), Cardiff, Wales: Alpha Omega.
- Jeffcoate, W.J., Lincoln, N.B., Selby, C., & Herbert M. (1986). Correlation between anxiety and serum prolactin in humans. *Journal of Psychosomatic Research*, **30** (2), 217-222.
- Johnson, S.G., Joplin, G.F., & Burrin, J.M. (1987). Direct assay for testosterone in saliva: relationship with a direct serum free testosterone assay. *Clinical Chemistry Acta.*, **163**, 309-318.
- Jones, G., Hanton, S., & Swain, A.B.J. (1994). Intensity and interpretation of anxiety symptoms in elite and non-elite sports performers. *Personality and Individual Differences*, **17**, 657-663.
- Jones, J., Murphy, E., Alaghband-Zadeh, J. (2004). *Modified method for extracted assay for female testosterone measurement on the Abbott Architect analyser*. Conference Proceedings. 23rd Joint Meeting of the British Endocrine Societies with the European Federation of Endocrine Societies, Brighton, UK.
- Kaplan, J.R., Mancuk, S.B., Clarkson, T.B., Lusso, F.M., & Taub, D.M. (1982). Social status, environment, and atherosclerosis in cynomolgus monkeys. *Arteriosclerosis*, **2**, 359-368.
- Kemeny, D.M. (1991). *A Practical Guide to ELISA*. Exeter: Pergamon Press.
- Kemeny, D.M. & Challacombe, S.J. (Eds.). (1991). *ELISA and other solid phase immunoassays: Theoretical and practical aspects*. Chichester: John Wiley & Sons.
- Kemeny, D.M. & Chantler, S. (1991). An introduction to ELISA. In D.M. Kemeny & S.J. Challacombe. (Eds.). *ELISA and other solid phase immunoassays: Theoretical and practical aspects*. Chichester: John Wiley & Sons.
- Kemper, T.D. (1990). *Social structure and testosterone: Explorations of the socio-bio-social chain*. New Brunswick: Rutgers University Press.
- Kemper, T.D. (1998). *Fantasy, females, sexuality, and testosterone* (commentary). *Behavioral and Brain Sciences*, **21**, 378-379.
- Keppel, G. & Wickens, T.D. (2002). Design and analysis: A researchers handbook (4th Ed., pp.393-395). NJ: Pearson Education Inc.
- Kirschbaum, C. & Hellhammer, D. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology*, **19**, 313-333.
- Kirschbaum, C., Kudielka, B.M., Gaab, J., Schommer, N.C., & Hellhammer, D.R. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine*, **61**, 154-162.

Kivlighan, K.T., Granger, D.A., Schwartz, E.B., Nelson, V., Curran, M., & Shirtcliff, E.A. (2004). Quantifying blood leakage into the oral mucosa and its effects on the measurement of cortisol, dehydroepiandrosterone, and testosterone in saliva. *Hormones and Behavior*, **46**, 39-46.

Koford, C. (1963). Rank of mothers and sons in bands of rhesus monkeys. *Science*, **141** (July), 356-357.

Kudielka, B.M., Schommer, N.C., Hellhammer, D.H., & Kirschbaum, C. (2003). Acute HPA axis response, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology*, **29**, 983-992.

Kuoppasalmi, K., Näveri, H., Rehunen, S., Härkönen, M., Adlercreutz, H. (1976) Effect of strenuous anaerobic running exercise on plasma growth hormone, cortisol, luteinizing hormone, testosterone, androstenedione, estrone and estradiol. *Journal of Steroid Biochemistry*, **7**, 823-829.

Labrie, F., Luu-The, V., Lin, S.X., et al. (2000). Intracrinology: role of the family of 17 β -hydroxysteroid, dehydrogenases in human physiology and disease. *Journal of Molecular Endocrinology*, **25**, 1-16.

Lashansky, G., Saenger, P., Fishman, K. et al. (1991). Normative data for adrenal steroidogenesis in a normal pediatric population. *Journal of Clinical Endocrinology and Metabolism*, **73**, 674-686.

Laughlin, G. & Barrett-Connor, E. (2001). Postmenopausal testosterone. *The Journal of Clinical Endocrinology and Metabolism*, **86** (4), Letter to Editor, 1834-1836.

Lincoln, G.A., Rowe, P.H., & Racey, R.A. (1974). *The circadian rhythm in plasma testosterone concentrations in man*. In J.Ashcroft, F.Ceresa, & F.Halberg. (Eds.). *Chronobiological Aspects of Endocrinology*. Stuttgart: FK Schattauer Verlag.

Lipson, S.F. & Ellison, P.T. (1989). Development of protocols for the application of salivary steroid analyses to field conditions. *American Journal of Human Biology*, **1**, 249-255.

Lobotsky, J., Wyss, H. I., Segre, E.J., & Lloyd, C.W. (1964). Plasma Testosterone in the Normal Woman. *Journal of Clinical Endocrinology*, **24**, 1261-1265.

Longcope, C. (1986). Adrenal and gonadal androgen secretion in normal females. *Journal of Clinical Endocrinology and Metabolism*, **15**, 213-227.

MacDougall, J.M., Dembroski, T.M., & Krantz, David, S. (1981). Effects of type of challenge on pressor and heart rate responses in type A and B women. *Psychophysiology*, **18** (1), 1-9.

Major, C.V., Read, S.E., Coates, R.A., Francis, A., McLaughlin, B.J., Millson, M., et al. (1991). Comparison of saliva and blood for human immunodeficiency virus prevalence testing. *Journal of Infectious Diseases*, **163**, 699-702.

- Mandel, I.D. (1993). Salivary diagnosis: Promises, promises. In D. Malamud, & L.A. Tabak. (Eds.). *Saliva as a diagnostic fluid. Annals of the New York Academy of Sciences*, **694** (pp. 1-10). New York, NY: New York Academy of Sciences.
- Mason, J.W. et al., (1973). Plasma cortisol and norepinephrine responses in anticipation of muscular exercise. *Psychosomatic Medicine*, **35**, 406-414.
- Massafra, C., De Felice, C., Agnusdei, D.P., Gioia, D., & Bagnoli, F. (1998). Androgens and Osteocalcin during the menstrual cycle. *Journal of Clinical Endocrinology and Metabolism*, **84** (3), 971-974.
- Mazur, A. (1973). A cross-species comparison of status in small established groups. *American Sociological Review*, **38** (5), 513-530.
- Mazur, A. (1976). Effects of testosterone on status in primate groups. *Folia Primatol*, **26**, 214-226.
- Mazur, A. (1985). A biosocial model of status in face-to-face primate groups. *Social Forces*, **64** (2), 377-402.
- Mazur, A. & Booth, A. (1998). Testosterone and dominance in men. *Behavioural and Brain Sciences*, **21**, 353-397.
- Mazur, A., Booth, A., & Dabbs Jr., J.M. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, **55** (1), 70-77.
- Mazur, A. & Lamb, T. (1980). Testosterone, status, and mood in human males. *Hormones and Behavior*, **14**, 236-246.
- Mazur, A. & Robertson, L. (1972). *Biology and social behavior*. New York: Free Press.
- Mazur, A., Susman, A.J., & Edelbrock, S. (1997). Sex difference in testosterone response to a video game contest. *Evolution and Human Behavior*, **18**, 317-326.
- McCaul, K.D., Gladue, B.A., & Joppa, M. (1992). Winning, losing, mood, and testosterone. *Hormones and Behavior*, **26**, 486-504.
- McGuire, M.T., Brammer, G.L., & Raleigh, M.J. (1986). Resting cortisol levels and the emergence of dominant status among male vervet monkeys. *Hormones and Behavior*, **20**, 106-117.
- McNair, D.M., Lorr, M., & Droppleman, L.F. (Eds.) (1971). *How to Use the Profile of Mood Sates (POMS) in Clinical Evaluations*. San Diego: Educational and Industrial Testing Services.
- Meikle, A.W., Stringham, J.D., Bishop, D.T., & West, D.W. (1988). Quantitating genetic and non genetic factors influencing androgen production and clearance rates in men. *Journal of Clinical Endocrinology and Metabolism*, **67**, 104-109.

Miller, W. & Tyrrell, J. (1995). *The adrenal cortex*. In P. Felig, J. Baxter., & Froham, L. (Eds.). *Endocrinology and Metabolism*, (3rd ed., pp. 555-712). New York: McGraw Hill.

Miner, V. & Longino, H. (1987). (Eds.). *Competition: A feminist taboo?* New York: Feminst Press.

Monaghan, E.P. & Glickman S.E. (1993). *Hormones and aggressive behavior*. In J.B. Becker, S.M. Breedlove, & D Crews. (Eds.). *Behavioral Endocrinology* (pp.261-285). Cambridge, MA: The MIT Press.

Mooradian, A.D., Morely, J.E., & Korenman, S.G. (1987). Biological actions of androgens. *Endocrine Review*, **8**, 1-27.

Morris, N.M., Udry, J. R., Khan-Dawood, F., & Dawood, M.Y. (1987). Marital sex frequency and midcycle female testosterone. *Archives of Sexual Behavior*, **16** (1), 27-37.

Moskowitz, D.S. (1993). Dominance and friendliness: On the interaction of gender and situation. *Journal of Personality*, **61**, 387-409.

Moskowitz, D.S., Suh, E.S., & Desaulniers, J. (1994). Situational influences on gender differences in agency and communication. *Journal of Personality and Social Psychology*, **66**, 751-761.

Moyer, K.E. (1968). Kinds of aggression and their physiological basis. *Communications in behavioral biology*, **2**, 65-87.

Mudrack, P.E. (1993). Relationship between dominance and achievement among self-report measures. *Psychological Reports*, **73**, 971-977.

Muncer, S.M., Campbell, A., Jervis, V., & Lewis, R. (2002). Ladettes, social representations and aggression. *Sex Roles*, **44**, 16-26.

Navarro, M.A., Juan, L., & Bonnin, M.R. (1986). Salivary testosterone: Relationship to total and free testosterone in serum. *Clinical Chemistry*, **32**, 231-232.

Navazesh, M. (1993). Methods for collecting saliva. In D. Malamud, & L.A. Tabak. (Eds.). *Saliva as a diagnostic fluid. Annals of the New York Academy of Sciences*, **694** (pp. 72-77). New York, NY: New York Academy of Sciences.Saliva.

Nicolau, G.Y., Haus, E., Lakatua, D.J. et al. (1985). Circadian and circannual variations of FSH, LH, testosterone, dehydroepiandrosterone-sulfate (DHEA-S), amd 17-hydroxy progesterone (17 OH-PROG) in elderly men and women. *Roumanian Journal of Medicine*, **23**, 223-246.

Obminski, Z. & Stupnicki, R. (1997). Comparison of the testosterone-to-cortisol ratio values obtained from the hormonal assays in saliva and serum. *Journal Sports Medicine Physical Fitness*, **37**, 50-55.

- O'Carroll, R. & Bancroft, J. (1984). Testosterone therapy for low sexual interest and erectile dysfunction in men: a controlled study. *British Journal of Psychiatry* **145** 146-151.
- O'Sullivan, M.J., Bridges, J.W., & Marks, V. (1979). Enzyme immunoassay: a review. *Annals of Clinical Biochemistry*, **16**, 221-239.
- Parker, L.N. (1991). Andrenarche. *Endocrinology and Metabolism Clinics of North America*, **20**, 71-83.
- Passelergue, P. & Lac, G. (1999). Saliva cortisol, testosterone and T/C ratio variations during a wrestling competition and during the post-competitive recovery period. *International Journal of Sports Medicine*, **20**, 109-113.
- Pusey, A., Williams, I., & Goodall, J. (1997). The influence of dominance rank on the reproductive success of female chimpanzees. *Science*, **277**, 828-31.
- Quissell, D.O. (1993). Steroid hormone analysis in human saliva. In D. Malamud, & L.A. Tabak. (Eds.). *Saliva as a diagnostic fluid. Annals of the New York Academy of Sciences*, **694** (pp. 143-145). New York, NY: New York Academy of Sciences.
- Raleigh, M.J., McGuire, M.T., Brammer, G.L., Pollack, D.B. & Yuwiler, A. (1991) Serotonergic mechanisms promote dominance acquisition in adult male vervet monkeys. *Brain Research*, **559**, 181-190.
- Read, G.F. (2001). Status report on measurement of salivary estrogens and androgens, Sabre Sciences, www.sabresciences.com/salivadiagnosticfluid.
- Read, G.F. (1989). *Hormones in saliva*. Tenovuo, J.O. (Eds.). Human Saliva: Clinical Chemistry and Microbiology. Boca Raton, FL: Press Inc.
- Read G.F. & Walker R. (1982). *Variation of salivary testosterone with age in men: discussion*. In G.F. Read, D. Riad-Fahmy, D., Walker, R.F., & Griffiths, K. (Eds.). Ninth Tenovus Workshop: Immunoassays of Steroids in Saliva. (pp.215-220), Cardiff, Wales: Alpha Omega.
- Rejeski, W.J., Brubaker, P.H., Herb, R.A., Kaplan, J.R., & Koritnik, D.R. (1988). Anabolic steroids and aggressive behaviour in cynomolgus monkeys *Journal of Behavioral Medicine*, **11** (1), 95-105.
- Rejeski, W.J., Gagne, M., Parker, P.E., & Koritnik, D.R. (1989). Acute stress reactivity from contested dominance in dominant and submissive males. *Behavioral Medicine*, **15**, 118-124.
- Rejeski, W.J., Parker, P.E., Gagne, M., & Koritnik, D.R. (1990). Cardiovascular and testosterone responses to contested dominance in women. *Health Psychology*, **9** (1), 35-47

- Remer, T., Pietrzik, K., & Manz, F. (1998). Short-term impact of a lactovegetarian diet on adrenocortical activity and adrenal androgens. *The Journal of Clinical Endocrinology and Metabolism*, **83** (6), 2123-2137.
- Renfrew, J. (1997). *Aggression and its causes. A biopsychosocial approach*. Oxford: Oxford University Press.
- Riad-Fahmy, D., Read, G.F., Walker, R.F., & Griffiths K. (1982). Steroids in saliva for assessing endocrine function. *Endocrine Reviews*, **3** (4), 367-395.
- Riad-Fahmy, D., Read, G.F., & Walker, R.F. (1983). Salivary steroid assays for assessing variation in endocrine activity. *Journal of Steroid Biochemistry*, **19**, 265-272.
- Richard, A.F. & Schulman, S.R. (1982). Sociobiology: Primate field studies. *Annual Review of Anthropology*, **11**, 231-255.
- Ridgeway, C.L. & Berger, J. (1986). Expectations, legitimation and dominance behaviour in task groups. *American Sociological Review* **51**, 603-617.
- Ridgeway, C.L., Berger, J., & Smith L. (1985). Nonverbal cues and status: An expectation states approach. *American Journal of Sociology*, **90**, 955-78.
- Ridley, M. (2003). *Nature via nurture: Genes, experience and what makes us human*. London: Fourth Estate.
- Rogers, L. (1999). *Sexing the brain*. London: Weidenfield & Nicolson.
- Rose, R.M., Gordon, T.P., & Bernstein, I.S. (1972). Plasma testosterone levels in the male rhesus: Influences of sexual and social stimuli. *Science, New Series*, **178** (4061), 643-645.
- Rosner, W. (1997). Errors in the measurement of free testosterone [letter]. *Journal of Clinical Endocrinology and Metabolism*, **82**, 2014-2015.
- Rosner, W. (2001). An extraordinarily inaccurate assay for free testosterone is still with us [Letter]. *Journal of Clinical Endocrinology and Metabolism*, **86** (6), 2903.
- Rowell, T.E. (1973). Social organization of wild talapoin monkeys. *American Journal of Physical Anthropology*, **38**, 593-598.
- Rushall, B.S. & Lippman, L.G. (1997). The role of imagery in physical performance. *International Journal for Sport Psychology*, **29**, 57-72.
- Salvador, A., Simón, V., Suay, F., & Llorens, L. (1987). Testosterone and cortisol responses to competitive fighting in human males: A pilot study. *Aggressive Behavior*, **13**, 9-13.
- Salvador, A., Suay, F., Gonzalez-Bono, E., & Serrano, M.A. (2003). Anticipatory cortisol, testosterone and psychological responses to judo competition in young men. *Psychoneuroendocrinology*, **28**, 364-375.

- Sapolsky, R.M. (1987). Stress, social status, and reproductive physiology in free-living baboons. In D. Crew (Ed.). *Psychobiology of reproductive behaviour: An evolutionary perspective*. Englewood Cliffs, NJ: Prentice Hall.
- Sapolsky, R.M. (1992). *Neuroendocrinology of the stress response*. In J. Becker, S. Breedlove, & D. Crew (Eds.). *Behavioral Endocrinology* (pp. 287-324). Cambridge, MA: MIT Press.
- Sapolsky, R.M., (1993). *The physiology of dominance in stable versus unstable social hierarchies*. In W.A. Mason & S.P. Mendoza (Eds.), *Primate social conflict* (pp. 171-204). New York: State University of New York Press.
- Sapolsky, R.M. (1998). *Why zebras don't get ulcers: An updated guide to stress, stress related diseases, and coping*. New York: W.H. Freeman and Company.
- Schmidt Mast, M. (2001). Gender differences and similarities in dominance hierarchies in same-sex groups based on speaking time. *Sex Roles*, **44**, 537-556
- Schürmeyer, T. & Nieschlag, E. (1982). *Salivary and serum testosterone under physiological and pharmacological conditions*. Ninth Tenovus Workshop: Immunoassays of Steroids in Saliva (pp.202-209). Cardiff, Wales: Alpha Omega Publishing.
- Schultheiss, O.C., Campbell, K.L., & McClelland, D.C. (1999). Implicit power motivation moderates men's testosterone responses to imagined and real dominance success. *Hormones and Behavior*, **36**, 234-241.
- Shimada, M., Takahashi, K., Ohkawa, T., Segawa, M., & Higurashi, M. (1995). Determination of salivary cortisol by ELISA and its application to the assessment of the circadian rhythm in children. *Hormone Research*, **44**, 213-217.
- Shirtcliff, E.A., Granger, D.A., & Likos, A. (2002). Gender differences in the validity of testosterone measured in saliva by immunoassay. *Hormones and Behavior*, **42**, 62-69.
- Shirtcliff, E.A., Granger, D.A., Schwartz, E., & Curran, M.J. (2001). Use of salivary biomarkers in biobehavioral research: Cotton-based sample collection methods can interfere with salivary immunoassay results. *Psycho-neuroendocrinology*, **26**, 165-173.
- Simpson, R.E. (2002). Aromatisation of androgens in women: Current concepts and findings. *Fertility and Sterility*, **77** (4), Suppl. 4, S6-S10.
- Sinha-Hikim, I., Arver, S., Beall, G., Shen, R. et al., (1998). The use of a sensitive equilibrium dialysis method for the measurement of free testosterone levels in healthy, cycling women and in the human immunodeficiency virus infected women. *Journal of Clinical Endocrinology and Metabolism*, **83** (4), 1312-1318.
- Smals, A.G.H., Kloppenbort, P.W.C., & Benraad, T.H. (1976). Cicannual cycle in plasma testosterone in man. *Journal of Clinical Endocrinology and Metabolism*, **42**, 979-982.

- Smith, R.G., Besch, P.K., Dill, P., & Buttram, V.C. (1979). Saliva as a matrix for measuring free androgens: Comparison with serum androgens in polycystic ovarian disease. *Fertility and Sterility*, **31**, 513.
- Smyth, J.M., Ockenfels, M.C., Gorin, A.A., Catley, D., Porter, L.S., Kirschbaum, C., Hellhammer, D.H., & Stone, A.A. (1997). Individual differences in the diurnal cycle of cortisol. *Psychoneuroendocrinology*, **22** (2), 89-105.
- Smyth, J.M., Ockenfels, M.C., Porter, L., Kirschbaum, C., Hellhammer, D.H., & Stone, A.A. (1998). Stressors and mood measured on a momentary basis are associated with salivary cortisol secretion. *Psychoneuroendocrinology*, **23** (4), 353-370.
- Somboonporn, W. & Davis, S.R. (2004). Testosterone effects on the breast: Implications for testosterone therapy for women. *Endocrine Reviews*, **25** (3), 374-388.
- Sternberg, R.J. (1993). *The Psychologists Companion* (3rd edition). Cambridge University Press.
- Stone, A.A., Schwartz, J.E., Smyth, J. Kirschbaum, C., Cohen, S., Hellhammer, D., & Grossman, S. (2001). Individual differences in the diurnal cycle of salivary free cortisol: a replication of flattened cycles for some individuals. *Psychoneuroendocrinology*, **26**, 295-303.
- Suay, F., Salvador, A., Gonzalez-Bono, E., Sanchis, C., Martinez, M., Martinez-Sanchis, S., Simon, V.M., & Montoro, J.B. (1999). Effects of competition and its outcome on serum testosterone, cortisol and prolactin. *Psychoneuroendocrinology*, **24**, 551-566.
- Sutton, J. R. & Casey, J.H. (1975). The adrenocortical response to competitive athletics in veteran athletes. *Journal of Clinical Endocrinology and Metabolism*, **40** (1), 135-138.
- Sutton, J.R., Coleman, M.J., Casey, J.H., & Lazarus, L. (1973). Androgen responses during physical exercise. *British Medical Journal*, **163**, 520-523.
- Svare, B.B. & Mann, M.A. (1983). *Hormonal influences on maternal aggression*. In B.B. Svare (Ed.), *Hormones and Aggressive Behavior* (pp. 91-104). New York: Plenum.
- Swift, A.D. & James, V.H.T. (1998). Discussion. Ninth Tenovus Workshop-Immunoassays of Steroids in Saliva (pp.215-220), Cardiff, Wales: Alpha Omega Publishing.
- Swift, A.D. (1989). *Salivary testosterone in the male adolescent*. In D.B. Ferguson (Ed.), *Steroid hormones in saliva* (pp. 69-79). Basel: Karger.
- Taieb, J., Benattar, C., Birr, A.S., & Lindenbaum, A. (2002). Limitations of Steroid Determination by Direct Immunoassay. *Clinical Chemistry*, **48**, 583-585.

Tames, F.J. & Swift, A.D. (1983). *The measurement of salivary testosterone*. In W.M. Hunter, & J.E.T. Corrie, (Eds). *Immunoassays for clinical chemistry*. New York: Churchill Livingstone.

Taylor, S.E., Klein, L.C., Brian P. et al. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, **107** (3), 411-429.

Thuma, J.R., Gilders, R., Verdun, M., & Loucks, A.B. (1995). Circadian rhythm of cortisol confounds cortisol responses to exercise: Implications for future research. *The American Physiological Society*.

Tinbergen, N. (1953). *Social behaviour in animals*. London: Methuen.

Tracy, L. (1991). *The secret between us: Competition among women*. Boston: Little Brown,

Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, **46**, 35-57.

Trivers, R.L. (1972). *Parental investment and sexual selection*. In B. Campbell (Ed.). *Sexual selection and the descent of man*. Chicago: Aldine.

Turkes, A.O., Turkes, A., Joyce, B.G., & Riad-Fahmy, D. (1980). A sensitive enzymeimmunosassay with a flourimetric end-point for the determination of testosterone in female plasma and saliva. *Steroids*, **35**, 89.

Udry, J.R. & Morris, N.M. (1977). The distribution of events in the human menstrual cycle. *Journal of Reproductive Fertility*, **51**, 419-425.

Valette, A., Seradour, B., & Boyer, J. (1975). Plasma testosterone levels during the menstrual cycle. *Journal of Clinical Endocrinology and Metabolism*, **40**, 160-161.

van den Berghe, P.L. (1974). Bringing beasts back in. *American Sociological Review* **39**, 777-788.

van Hoof, J.A. & van Schaik, C.P. (1992). *Cooperation in competition: The ecology of primate bonds*. In A. Harcourt & F.B.M. de Waal (Eds.). *Coalitions and alliances in humans and other animals*. Oxford: Oxford University Press.

Veldhuis, J.D., King, J.C., Urban, R.J. et al. (1987). Operating characteristics of the male hypothalamo-pituitary-axis: Pulsatile release of testosterone and follicle-stimulating hormone and their temporal coupling with leutinizing hormone. *Journal of Clinical Endocrinology and Metabolism*, **65**, 929-841.

Vermeulen, A. (1998). Plasma androgens in women. *Journal of Reproductive Medicine*, **43**, 725-33.

Vermeulen, A. & Verdonck, L. (1972). Some studies on the biological significance of free testosterone. *Journal of Steroid Biochemistry*, **3**, 421-426.

- Vermeulen, A. & Verdonck, L. (1976). Plasma androgen levels during the menstrual cycle. *American Journal of Obstetrics and Gynaecology*, **125** (4), 491-494.
- Vermeulen, A., Verdonck, L., & Kaufman, J.M. (1999). A critical evaluation of simple methods for the estimation of free testosterone in serum. *Journal of Clinical Endocrinology and Metabolism*, **84** (10), 3666-3672.
- Vining, R.F., McGinley, R.A., & Symons, R.G. (1983). Hormones in saliva: Modes of entry and consequent implications for clinical interpretation. *Clinical Chemistry*, **29**, 1752-1756.
- Vittekk, J., L'Hommedieu, D.G., Gordon, G.G., Rappaport, S.C., & Southren, A.L. (1985). Direct radioimmunoassay (RIA) of salivary testosterone: Correlation with free and total serum testosterone. *Life Sciences*, **37**, 711-716.
- Vogel, R.B., Books, C.A., Ketchum, C.A., Ketchum, C., Zauner, C.W., & Murray, F.T. (1985). Increase of free and total testosterone during submaximal exercise in normal males. *Medicine and Science in Sports and Exercise*, **17**, 119-123.
- Walker, R.F., Wilson, D.W., Read, G.F., & Riad-Fahmy, D. (1980). Assessment of testicular function by the radioimmunoassay of testosterone in saliva. *International Journal of Andrology*, **3**, 105.
- Wang, C., Plymate, S., Nieschlag, E., & Paulsen, C.A. (1981). Salivary testosterone in men: Further evidence of a direct correlation with free serum testosterone. *Journal of Clinical Endocrinology and Metabolism*, **53** (5), 1021-1024.
- Wickings, E.J. & Dixon, A.F. (1992). Testicular function secondary sexual development and social status in male mandrills (*mandrillus sphinx*). *Physiology and Behavior*, **52**, 909-916.
- Weisfeld, G. (1994). *Aggression and dominance in the social world of boys*. In J. Archer (Ed.). *Male Violence* (pp310-331). London: Routledge.
- Wilson, E.O. (1975). *Sociobiology: The new synthesis*. Cambridge, MA: Harvard University Press.
- Winslow, J.T., Ellingboe, J., & Miczek, K.A. (1988). Effects of alcohol on aggressive behaviour in squirrel monkeys: Influence of testosterone and social context. *Psychopharmacology*, **95**, 36-363.
- Xinari, C. (2004). *Bare essentials: Gender fictions, embodiment matters*. Unpublished Doctoral Dissertation, University of Edinburgh, Edinburgh: Scotland.
- Zitmann, M. & Nieschlag, E. (2001). Testosterone levels in healthy men and the relation to behavioural and physical characteristics: Facts and constructs. *European Journal of Endocrinology*, **144**, 183-197.

Zumoff, B., Strain, G.W., Miller, L.K., & Rosser, W. (1995). 24 hour mean plasma testosterone concentration declines with age in normal premenopausal women. *Journal of Clinical Endocrinology and Metabolism*, **80** (4), 1429-1430.

APPENDICES

Chemical Reagents

- Wash buffer: prepared by the addition of 0.05% Tween 20 to phosphate buffered saline (PBS) (pH 7.4)
- Assay buffer: prepared by the addition of 0.1% BSA to PBS
- Blocking buffer: prepared by the addition of 0.1% BSA to PBS
- Substrate buffer: sodium acetate citric acid buffer (pH 4.2)
- Substrate solution: 12mL of substrate buffer was mixed with 300µl of substrate stock (tetramethylbenzadine at 10mg/mL) and 40µl of hydrogen peroxidase
- De-ionised water of greater than 15 ohms is used to prepare the wash solution
- Testosterone Standards: Concentrations of 0, 0.05, 0.25, 1, 5, 10µl of testosterone were prepared in assay buffer from a stock testosterone solution of 200ng/mL.

Consumables

- 1.5ml, 39 x 10 mm Micro tubes for aliquoting
- Fisherbrand blue pipette tips 200-1000µl
- Greiner bio-one medium bind plates (96 wells)



School of Philosophy, Psychology, and Language Sciences

Testosterone Circadian Activity in Women

Exclusion Criteria

This study measures female salivary testosterone and cortisol. The levels of testosterone in women are extremely low (parts per billion) and the measurement techniques used are extremely sensitive to a variety of interfering factors. Consequently, we are required to pay particular attention to exclusion criteria in the study. The following statements provide a preliminary but comprehensive guide to eligibility and are intended to assist both you and ourselves in determining who would be suitable to participate.

You will **NOT** be eligible to participate if:

1. Your menstrual cycles are irregular:
Even women who report regular cycles actually experience cycles of differing lengths. Consequently, our definition of regular cycles is taken to mean that your previous six cycles have been between 24 and 35 days in duration
2. You are currently taking the birth control pill or have stopped taking the pill within the past 90 days
3. You are administering any other form of hormonal treatment
4. You are on a strict diet or are seriously restricting your calorific intake
5. You have a previous history of kidney disease
6. You have a previous history of liver disease
7. You have pronounced acne
8. You have hirsutism combined with irregular menstrual cycles
9. You think you might be considered overweight
10. You have consumed illicit drugs during the previous six months



School of Philosophy, Psychology, and Language Sciences

Establishing Normal Ranges and Circadian Rhythm Profiles for Salivary Testosterone In Healthy Adult Females

Declaration of Informed Consent

I give my consent to participate in this study that investigates the extent to which the hormones testosterone and cortisol fluctuate throughout the course of a day. I consent to publication of study results so long as the information is anonymous and disguised so that no identification can be made. I further understand that although a record will be kept of my having participated in the study, a number only will identify all study data collected from my participation.

1. I have been informed that my participation in this study will involve me collecting multiple saliva samples into a sterile collection container and that saliva production will be stimulated by chewing on a small amount of sugar free gum
2. I have been informed that the general purpose of this study is to determine the extent to which the hormones testosterone and cortisol fluctuate throughout the course of a single day. Moreover, this data will be used to determine normal ranges for free testosterone levels in women of varying age
3. I have been informed that there are no known expected discomfort or risks involved with my participation in this study
4. I have been informed that the experimenter will courteously answer any questions regarding the procedures in this study
5. I have been informed I will receive payment of £20 for successful completion of the study
6. I have been informed that I am free to withdraw from the study at any time

Concerns about any aspect of this study may be referred to the Principal Investigator: Martin Sharp, Dept of Psychology, University of Edinburgh, EH8 9JZ; phone... ..Email: martin.sharp@ed.ac.uk

Experimenter: Signature _____

Participant: Signature _____

Participant: Print Name _____

Date: _____



Diurnal Variation of Salivary Testosterone and Cortisol in Human Females

Thank you for agreeing to participate in this study that investigates the fluctuation of salivary hormones throughout the course of a day. The study will involve you collecting a total of 8 small saliva samples. The following sheets contain all the information you require to participate in this study. Please accurately follow the step-by-step procedures detailed below.

Conditions of Collection

On the days that you collect saliva there are two conditions you will need to adhere to. Your testosterone levels are affected, in an unpredictable manner, by both exercise and alcohol. Consequently, it is *extremely* important that you do not exercise or drink alcohol in the 24 hours before collection or during collection of saliva samples themselves. Exercise is taken to mean any organised physical activity such as swimming, cycling, yoga or going to the gym.

Collection of Samples

Each time you collect saliva the procedure should be identical. The steps you should follow are outlined below. In order to prevent tiny cuts or micro-abrasions to the mouth please do not eat or brush your teeth for approximately *90 minutes* prior to saliva collection.

- Step 1. Rinse and swill mouth out thoroughly three times with tap/filter water
- Step 2. Wait three minutes for the environment in your mouth to normalise
- Step 3. Chew on quarter of a stick of the sugar free gum provided. This will aid you in producing saliva and the step *must* be included
- Step 4. Whilst retaining the gum in your mouth, spit away the first mouthful of saliva. This will get rid of unwanted cellular elements in the gum
- Step 5. Continue chewing on gum and spitting in the tube until you have deposited 5ml of saliva into collection container. Please then ensure caps are replaced tightly

Overleaf you will see a timetable for collecting the samples. Although it is not a problem if you begin collection 5 minutes either side of the specified time please try and get as close to them as possible and mark down the time you collected the sample. If any event occurs in the two hours *prior* to each collection, which stresses or upsets you, which makes you especially happy or which might in any way be considered out of the ordinary then please make a

note of it in the right hand comments column. This might include events such as unavoidable physical exercise, arguments, stress and sexual activity.

Finally, please pay special attention to the labelling on the tubes. Each tube will be marked with the appropriate collection time (i.e. 9am, 11pm etc.) and the tubes will be in plastic bags marked with the appropriate day of your menstrual cycle. If you collect a sample in the wrong tube don't worry; just re-mark the tube with the correct time and let the experimenter know when handing the samples back.

Storage of Samples

Because you may wish to collect samples whilst out and about you will need to give a little thought to how the samples are stored. Once you have collected a sample (for example 11am) you can do one of two things with it. Either, place it in a fridge immediately, or keep it with you in a cool place (preferably below room temperature). If you are going to choose this option then please keep the sample out of direct sunlight and don't let it get too hot (radiators are definitely to be avoided). Once you have collected all the samples for a particular day they will need to be placed in a fridge overnight or until you are ready to return them.

Handing Samples Back

Please call Martin Sharp on (M) to arrange a suitable time for sample collection/delivery. Should you experience any difficulties with the procedure then please do not hesitate to call me on the number above.

Payment

You will be paid, either in cash or by cheque, for your participation once all samples have been returned.

Menstrual Cycle Dates

1st day of cycle

Last day of cycle

*Collection Schedule**DAY 1*

	Suggested collection time	Actual time collected	Comments
Sample 1	9am		
Sample 2	11am		
Sample 3	1pm		
Sample 4	3pm		
Sample 5	5pm		
Sample 6	7pm		
Sample 7	9pm		
Sample 8	11pm		



Declaration of Informed Consent

Endocrine Response to Non-Physical Competition in Females

I give my consent to participate in this study which investigates potential modification in the levels of salivary hormonal markers in relation to non-physical competition. I consent to publication of study results so long as the information is anonymous and disguised so that no identification can be made. I further understand that although a record will be kept of my having participated in the study, a number only will identify all data collected from my participation.

1. I have been informed that participation in this study will involve me depositing multiple saliva samples into universal collection containers and completing questionnaires on mood, anxiety, and competition
2. I have been informed that the general purpose of this study is to determine how hormones change in relation to non-physical competition
3. I have been informed that there are no known expected discomfort or risks involved with my participation in this study
4. I have been informed that the experimenter will gladly answer any questions regarding the procedures in this study at any stage
5. I have been informed that I am free to withdraw from the study at any time

Experimenter: _____

Name of Participant (*please print*): _____

Signature of Participant: _____

Date: _____



Endocrine Response to Female Non-Physical Competition

Dear

Thank you for agreeing to participate in this research programme. This study is designed to examine whether or not females experience a change in the level of circulating salivary hormones in relation to competition. You will be asked to collect saliva samples at pre-determined times both prior to and following competition. In order to determine a baseline measurement you will also need to collect saliva on a non-competition day 48 hours prior to competition. In addition, you will be required to complete questionnaires on mood, anxiety, and competition

Detailed instructions, including the times and days on which you should collect samples, appear on the following pages. For each collection event please accurately follow the procedure detailed as this ensures that all samples have been collected under the same conditions.

Should you experience any difficulties with the procedure or require additional information then please do not hesitate to contact the principal investigator:

Martin Sharp
email: Martin.Sharp@ed.ac.uk
phone: 0131 650 3424 (W)
..... (M)

General Instruction to Participants when Collecting Saliva Samples

Please ensure these steps are followed every time you collect saliva

- Step 1. Rinse mouth *thoroughly* three times with tap/filter water — this removes trace elements of blood which can contaminate the sample
 - Step 2. Wait one minute for environment in mouth to normalise
 - Step 3. Start chewing on a quarter stick of sugar free gum (this is to stimulate saliva production which should only take a few moments)
 - Step 4. Whilst retaining the gum in your mouth spit the first mouthful of saliva away (this gets rid of unwanted debris from the gum)
 - Step 5. Continue chewing on the gum
 - Step 6. All additional saliva should be deposited into the universal collection container provided. Continue chewing on the gum and collecting saliva until you have deposited saliva upto the 5mL mark (actual saliva and not foam)
 - Step 7. Return all collection tubes to investigator as soon as possible. If at all possible all containers should be stored in a fridge until their return
- Note:* Please do not brush your teeth or eat in the twenty minutes prior to any saliva sampling as this can lead to trace elements of blood in the mouth rendering the samples useless

BASELINE — 48hrs PRIOR TO COMPETITION

Baseline Collection on Non-Competition Days

1. These baseline measurements should be taken 48 hours prior to competition. E.g., if your competition is on a Friday then these measurements should be taken on the preceding Wednesday
2. To work out sample collection times determine when your competition is likely to start (e.g. 3pm) and then count three hours back. This will give you the time when you should collect your first sample. Subsequent samples should be collected every hour. Samples five and six should be collected at 30 min intervals
3. Please follow the saliva collection procedure detailed earlier
4. Please ensure that the number, day and time on the collection container correspond to the collection event — you will be supplied with a bag marked BASELINE
5. Please return the collection tubes to the investigator as soon as possible. Samples should be refrigerated until their return
6. Please ensure the questions about your attitude to the competition (overleaf) are completed
7. Please complete the questionnaires overleaf at the times indicated in the table
8. Please do not exercise, engage in sexual activity, or consume alcohol prior to or during the saliva collection

Sample One	Sample Two	Sample Three	Sample Four	Sample Five	Sample Six
9am	3 hours prior to comp	2 hours prior to comp	1 hr prior to comp	30 mins prior to comp	Immediately prior to Start of Comp
				Complete CSAI-2	
			Complete POMS		

Question for Baseline Collection on Non-Competition Days

Question 1. Have you drunk alcohol at any point during the previous twenty-four hours (if yes, how much)?

Question 2. Have you engaged in any sexual activity during the collection times?

Question 3. Have you fasted during the course of then day?

Question 4. Has your diet changed from that of a normal day?

DAY OF COMPETITION

Collection Sheet for Day of Competition

- 1. The table below indicates when you should collect the saliva samples and when you should complete the questionnaires. In addition you should complete the questions overleaf at any time you wish during the day
- 2. Please follow the previous instruction on for collecting saliva — page 2
- 3. Once you have collected each sample please make a note of the time in the table below
- 4. Please ensure that the number on the labels matches the corresponding sample time.
- 5. Please return containers to investigator as soon as possible after final collection. If possible store them in a refrigerator until then.

Pre-Competition

	Sample One	Sample Two	Sample Three	Sample Four	Sample Five	Sample Six
	9am	3 hours prior to comp	2 hours prior to comp	1 hr prior to comp	30 mins prior to comp	Immediately prior to comp
Actual Time						
POMS				Complete POMS		
CSAI-2					Complete CSAI-2	
Questions						Complete questions

Collection Sheet for Day of Competition

Pre-Competition

	Not at all	A little	Moderately	Extremely
	1	2	3	4
1. How important is this competition to you?				
2. How would you rate your interest in winning this competition?				
3. How would you rate your chances of winning?				

Question 4. Have you followed any form of visualisation process or cognitive rehearsal?

Question 5. How would you rate your opponent's skill?

- a) much worse than mine
- b) a bit worse than mine
- c) about the same
- d) a bit better
- e) a lot better than me
- f) no idea

Collection Sheet for Day of Competition

Post-Competition

	Sample One	Sample Two	Sample Three	Sample Four	Sample Five
	Immediately end of Comp	30 mins post comp	1 hour post comp	2 hours post comp	3 hours post comp
			Complete POMS		
				Questions	
Time					

Post-Match Questions

1. Did you win or lose? _____

2. Regardless of outcome do you feel your performance was:
 - a) a lot better than expected
 - b) somewhat better
 - c) about the same
 - d) not quite as good
 - e) a lot worse than I expected to do

3. Estimate the challenge or your opponent?
 - a) they were more of a struggle than I thought
 - b) they performed at the level I expected
 - c) they were not as challenging as I had thought they would be

4. How much do you feel you contributed to the outcome of the match?
 - a) not very much at all
 - b) a little bit
 - c) quite a lot
 - d) a lot

5. How close do you think the outcome of the match was?
 - a) not very
 - b) reasonably
 - c) very